

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 19:08:57 ; Search time 61 Seconds
(without alignments)
668.438 Million cell updates/sec

Title: US-09-308-397-2

Perfect score: 1518

Sequence: 1 MTKTAFLEAGGAGYLGWGR.....QTAHLAHEVDQASIVALLEK 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1518	100.0	306	19	FabD polypeptide s
2	1518	100.0	306	19	CPE 34 protein seq
3	1515	99.8	306	19	S. pneumoniae prot
4	1514	99.7	306	22	Streptococcus pneu
5	1048	69.0	308	23	Lactococcus lactis
6	1036	68.2	314	23	Streptococcus poly
7	996.5	65.6	308	23	Streptococcus poly
8	937	61.7	313	22	Enterococcus faeca
9	863	56.9	182	19	Carboxy terminal r
10	670	44.1	137	18	S. pneumoniae malo

11	643.5	42.4	309	22	AAU34532	E. coli cellular p
12	636	41.9	312	22	AAU3382	Haemophilus influe
13	613	40.4	122	19	AAW60857	Amino terminal reg
14	588.5	38.8	313	23	ABB48893	Listeria monocytog
15	585	38.5	130	19	AAW61018	Streptococcus pneu
16	556.5	36.7	302	22	AAU33964	Staphylococcus aur
17	556.5	36.7	311	22	AAU36537	Staphylococcus aur
18	556.5	36.7	312	22	AAU36301	Pseudomonas aerugi
19	550.5	36.3	308	21	AAU70134	Staphylococcus aur
20	547.5	36.1	308	21	AAW60032	Amino acid sequenc
21	547.5	36.1	312	22	AAE02192	Staphylococcus aur
22	547.5	36.1	332	22	AAE02191	Staphylococcus aur
23	528	34.8	308	20	AAU34892	Chlamydia pneumoni
24	515.5	34.0	311	23	ABP39979	Staphylococcus epi
25	454.5	29.9	309	20	AAU36958	Protein involved i
26	439.5	29.0	327	21	AAU18524	Arabidopsis thalia
27	439.5	29.0	327	21	AAU18524	Arabidopsis thalia
28	439.5	29.0	367	21	AAU18523	Arabidopsis thalia
29	439.5	29.0	367	21	AAU18523	Arabidopsis thalia
30	439.5	29.0	393	21	AAU18522	Arabidopsis thalia
31	439.5	29.0	393	21	AAU18522	Arabidopsis thalia
32	433	28.5	330	21	AAU07678	Amino acid sequenc
33	419	27.6	3413	19	AAW52849	A. mediterranei xi
34	416.5	27.4	309	22	AAU35697	Helicobacter pylor
35	414.5	27.3	1402	22	AAU81115	Mycobacterium tube
36	414.5	27.3	1402	22	AAU81115	Protein encoded by
37	413.5	27.2	6797	22	AAU31558	Pimaricin biosynth
38	405	26.7	223	22	AAU82280	S. epidermidis ope
39	403	26.5	5435	22	AAE10145	Streptomyces nous
40	399.5	26.3	309	21	AAU47615	Streptomyces nous
41	399.5	26.3	349	21	AAU47615	Arabidopsis thalia
42	399.5	26.3	375	21	AAU47613	Arabidopsis thalia
43	399	26.3	405	21	AAU26155	B. cereus zwitterm
44	397.5	26.2	11096	22	AAE10129	Streptomyces nous
45	379.5	25.0	412	23	AAU77694	Actinomycete monen

ALIGNMENTS

RESULT 1

AAW60856
ID AAW60856 standard; Protein; 306 AA.

AC AAW60856;

DT 24-SEP-1998 (first entry)

DE FabD polypeptide sequence.

KW FabD protein; malonyl-CoA:ACP family; diagnosis; infection; vaccine; screen.

OS Streptococcus pneumoniae.

PN WO9822133-A1.

PD 28-MAY-1998.

PF 14-NOV-1997; 97WO-US20992.

PR 18-NOV-1996; 96US-0031160.

PA (SMK) SMITHKLINE BEECHAM CORP.

PI Gentry DR, Lonsdale JT, Payne DU, Pearson SC, Van Aller G;

DR WPI; 1998-312173/27.

DR N-PSDB; AAU37239.

PI New isolated Streptococcus pneumoniae FabD gene - used to develop products for the diagnosis, prevention and treatment of bacterial diseases, particularly S. pneumoniae infection

PT

XX PS Claim 2; Page 6; 45pp; English.

XX CC The present sequence represents a FabD protein of Streptococcus

CC CC pneumoniae 0105993 (NCIMB 40800). The novel FabD polypeptides are

CC CC related to other proteins of the malonyl-CoA:ACP family. The products

CC CC can be used for the diagnosis of Streptococcus pneumoniae infections.

CC CC Vectors containing the FabD DNA sequence can be administered directly

CC CC to a mammal to produce the FabD peptide to provoke an antibody/T-cell

CC CC response in order to prevent a disease. The peptide can be used to

CC CC screen for compounds which modulate its activity.

XX SQ Sequence 306 AA;

Query Match 100.0%; Score 1518; DB 19; Length 306;

Best Local Similarity 100.0%; Pred. No. 3.2e-131;

Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFAGGAGQAYLGMRDFYDQYPIVKETIDRASQVGLDYRLIDTEEDKLNQTRY 60

DB 1 MTKTAFAGGAGQAYLGMRDFYDQYPIVKETIDRASQVGLDYRLIDTEEDKLNQTRY 60

QY 61 TOPAILATSVAIYRLLEQKGYQPDVAGLSLGEYSALVAGSALDGEDAVALVAKRGAYME 120

DB 61 TOPAILATSVAIYRLLEQKGYQPDVAGLSLGEYSALVAGSALDGEDAVALVAKRGAYME 120

QY 121 EAAPADSGKMWAVLNPVEIEEACOKASELGWVTPANTPAQIVIAGEVAVDRAVEL 180

DB 121 EAAPADSGKMWAVLNPVEIEEACOKASELGWVTPANTPAQIVIAGEVAVDRAVEL 180

QY 181 LQEGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAVMQKEDI 240

DB 181 LQEGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAVMQKEDI 240

QY 241 AQLLTQVKPEVPFYESIGVMQEGAGISNFIEIGPGKVLGGFYKKIDQTAHLAHVEDQASL 300

DB 241 AQLLTQVKPEVPFYESIGVMQEGAGISNFIEIGPGKVLGGFYKKIDQTAHLAHVEDQASL 300

QY 301 VALLEK 306

DB 301 VALLEK 306

RESULT 2

AAW01031

ID AAW01031 standard; Protein; 306 AA.

AC AAW01031;

XX 02-OCT-2001 (first entry)

XX CFE 34 protein sequence.

XX Antibacterial; vaccine; gene therapy; bacterial cell wall viability;

XX CFE; CEG; Conserved Essential Gene; bacterial infection;

XX antisense therapy; antibiotic resistance.

OS Streptococcus pneumoniae.

XX WO200149721-A2.

XX 12-JUL-2001.

XX 29-DEC-2000; 2000WC-US35604.

XX 30-DEC-1999; 99US-0174089.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;

XX Thanassi JA;

XX WPI; 2001-496721/54.

DR N-PSDB; AAW90730.

XX Nucleic acids encoding conserved essential genes involved in bacterial

PT replication which are potential targets for the treatment of antibiotic

PT resistant bacterial infections -

XX Claim 27; Pages 272-273; 380pp; English.

XX The present invention relates to nucleic acids (AAW90701-AAW90918)

CC encoding polypeptides (AAW01002-AAW0114), which are essential for the

CC viability of a bacterial cell wall. The acronym CFE stands for "CEG For

CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic

CC acids are useful for detecting the presence of proteins essential for the

CC viability of a bacterial cell wall in samples such as cells, tissues,

CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,

CC and for detecting corresponding target nucleic acid molecules with

CC complementary sequences. The nucleic acids are also useful for

CC determining whether a genomic nucleotide sequence of interest is

CC essential for viability of a bacterial cell or whether it resides within

CC an operon, by integrating an exogenous nucleotide sequence comprising a

CC portion of an open reading frame of the genomic sequence of interest

CC (comprising 200-500 base pairs) into the genomic sequence of interest

CC which confers a selectable phenotype to the cell, and determining cell

CC viability with a selection agent such as chloramphenicol. The nucleic

CC acids and proteins are also useful as vaccines and for treating bacterial

CC infections with gene therapy and antisense therapy. The nucleic acids

CC also enable identification of targets suitable for the treatment of

CC antibiotic resistant bacterial infections.

XX SQ Sequence 306 AA;

Query Match 100.0%; Score 1518; DB 22; Length 306;

Best Local Similarity 100.0%; Pred. No. 3.2e-131;

Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFAGGAGQAYLGMRDFYDQYPIVKETIDRASQVGLDYRLIDTEEDKLNQTRY 60

DB 1 MTKTAFAGGAGQAYLGMRDFYDQYPIVKETIDRASQVGLDYRLIDTEEDKLNQTRY 60

QY 61 TOPAILATSVAIYRLLEQKGYQPDVAGLSLGEYSALVAGSALDGEDAVALVAKRGAYME 120

DB 61 TOPAILATSVAIYRLLEQKGYQPDVAGLSLGEYSALVAGSALDGEDAVALVAKRGAYME 120

QY 121 EAAPADSGKMWAVLNPVEIEEACOKASELGWVTPANTPAQIVIAGEVAVDRAVEL 180

DB 121 EAAPADSGKMWAVLNPVEIEEACOKASELGWVTPANTPAQIVIAGEVAVDRAVEL 180

QY 181 LQEGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAVMQKEDI 240

DB 181 LQEGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAVMQKEDI 240

QY 241 AQLLTQVKPEVPFYESIGVMQEGAGISNFIEIGPGKVLGGFYKKIDQTAHLAHVEDQASL 300

DB 241 AQLLTQVKPEVPFYESIGVMQEGAGISNFIEIGPGKVLGGFYKKIDQTAHLAHVEDQASL 300

QY 301 VALLEK 306

DB 301 VALLEK 306

RESULT 3

AAW80671

ID AAW80671 standard; Protein; 306 AA.

AC AAW80671;

XX 24-DEC-1998 (first entry)

XX S. pneumoniae protein (ACPtrans).

XX Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;

XX virulence; antibody; infection; detection; treatment; hypochetical;

XX cell wall biosynthetic; external target; minimal gene set protein.

XX Streptococcus pneumoniae.
 OS WO9826072-A1.
 XX 18-JUN-1998.
 XX 09-DEC-1997; 97WO-US22578.
 XX 13-DEC-1996; 96US-0036281.
 XX (ELIL) LILLY & CO ELI.
 XX Baltz RH, Burgert SG, Dehoff BS, Hoskins JA, Jaskunas SR;
 PI Mills BJ, Norris FH, Peery RB, Rostek PK, Rostek PR;
 PI Skatrud PL, Smith MC, Solenberg PU, Treadway PU;
 PI Young Bellido ML;
 XX WPI; 1998-348529/30.
 XX Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
 PT for evaluating gene expression, and identification of virulence
 PT genes
 XX Claim 3; Pages 259-259; 333pp; English.
 XX This sequence represents a Streptococcus pneumoniae protein. The
 CC invention provides DNA sequences (AAV655201 to AAV655304) from the
 CC Streptococcus pneumoniae genome and corresponding protein sequences
 CC (AAW80605 to AAW80728). The protein sequences are classified as
 CC hypothetical, cell wall biosynthetic, external target, or minimal gene
 CC set proteins. A recombinant host containing a vector comprising any of
 CC the above nucleic acids can be used for the recombinant expression of the
 CC proteins. The invention also provides a DNA chip having arrayed on it at
 CC least 15 base pair fragment of any one or more of these DNA sequences.
 CC The DNA chip can be used methods for evaluating gene expression in S.
 CC pneumoniae and for identifying virulence genes in S. pneumoniae.
 CC Antibodies that selectively bind to the above proteins or peptide
 CC fragments can be used to treat S. pneumoniae infection. The antibodies
 CC can also be used to detect S. pneumoniae cells.
 XX SQ Sequence 306 AA;
 Query Match 99.8%; Score 1515; DB 19; Length 306;
 Best Local Similarity 99.7%; Pred. No. 6.1e-131;
 Matches 305; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTKTFLFAGGGAQYLGMRDFYDQYPIVKETIDRASQVGLGYDLRYLIDTEEDKLNTRY 60
 DB 1 MTKTFLFAGGGAQYLGMRDFYDQYPIVKETIDRASQVGLGYDLRYLIDTEEDKLNTRY 60
 QY 61 TOPAILATSVAIYRLLOEKGYQDMVAGLSLGEYSALVAGSALFDFAVALVAKRGAYME 120
 DB 61 TOPAILATSVAIYRLLOEKGYQDMVAGLSLGEYSALVAGSALFDFAVALVAKRGAYME 120
 QY 121 EAPADSGKMWAVLNPVEVIEEACOKASELGVNTPANNYTPAQIVLAGEVAVDRAVEL 180
 DB 121 EAPADSGKMWAVLNPVEVIEEACOKASELGVNTPANNYTPAQIVLAGEVAVDRAVEL 180
 QY 181 LOEAGAKELIPKVSQPHHTALLEPASQKLAETLAQVSFSDFTCLVGNTEAAVWQKEDI 240
 DB 181 LOEAGAKELIPKVSQPHHTALLEPASQKLAETLAQVSFSDFTCLVGNTEAAVWQKEDI 240
 QY 241 AQLLTRQVKEPVRVYESIGVMOEAGISNFIIEGPKVLSGFVKIKIDQTAHLAHVEDQASL 300
 DB 241 AQLLTRQVKEPVRVYESIGVMOEAGISNFIIEGPKVLSGFVKIKIDQTAHLAHVEDQASL 300
 QY 301 VALLEK 306
 DB 301 VALLEK 306
 RESULT 4

AAU37987
 ID AAU37987 standard; Protein; 306 AA.
 XX AC AAU37987;
 XX 14-FEB-2002 (first entry)
 XX DE Streptococcus pneumoniae cellular proliferation protein #416.
 XX KW Antisense; prokaryotic cellular proliferation protein;
 XX KW antibiotic; antibacterial; drug design.
 XX OS Streptococcus pneumoniae.
 XX WO200170955-A2.
 XX 27-SEP-2001.
 XX 21-MAR-2001; 2001WO-US09180.
 XX 21-MAR-2000; 2000US-191078P.
 XX 23-MAY-2000; 2000US-206848P.
 XX 26-MAY-2000; 2000US-207727P.
 XX 23-OCT-2000; 2000US-242578P.
 XX 27-NOV-2000; 2000US-253625P.
 XX 22-DEC-2000; 2000US-257931P.
 XX 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 PI WPI; 2001-611495/70.
 DR N-PSDB; AAG55846.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX Example 3; Seq ID No 13580; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 306 AA;
 Query Match 99.7%; Score 1514; DB 22; Length 306;
 Best Local Similarity 99.7%; Pred. No. 7.5e-131;
 Matches 305; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTKTFLFAGGGAQYLGMRDFYDQYPIVKETIDRASQVGLGYDLRYLIDTEEDKLNTRY 60
 DB 1 MTKTFLFAGGGAQYLGMRDFYDQYPIVKETIDRASQVGLGYDLRYLIDTEEDKLNTRY 60
 QY 61 TOPAILATSVAIYRLLOEKGYQDMVAGLSLGEYSALVAGSALFDFAVALVAKRGAYME 120

Db 61 TOPAILATSVAIYRLLOEKGYOPDMVAGLSGEYSALVAGSALDFEDAVAVAKRGAYME 120
 QY 121 EAAPADSGKWAVLNTPEVIEEACQKASELGVVTPANTPAQIVIAGEVAVDRAVEL 180
 Db 121 EAAPADSGKWAVLNTPEVIEEACQKASELGVVTPANTPAQIVIAGEVAVDRAVEL 180
 QY 181 LOEAGAKRLIPLKVSQPHHTALLPASPQKLAETLAQVFSDFTCPLVGNTEAAVMQKEDI 240
 Db 181 LOEAGAKRLIPLKVSQPHHTALLPASPQKLAETLAQVFSDFTCPLVGNTEAAVMQKEDI 240
 QY 241 AQLLTROVKEPVRVYESIGVMQEAQISNFIETIGPGKVLGSGFKKIDRTAHLAHVEDQASL 300
 Db 241 AQLLTROVKEPVRVYESIGVMQEAQISNFIETIGPGKVLGSGFKKIDRTAHLAHVEDQASL 300
 QY 301 VALLEK 306
 Db 301 VALLEK 306

RESULT 5
 ABB54086
 ID ABB54086 standard; Protein; 308 AA.
 AC ABB54086;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Lactococcus lactis protein fabd.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR2807446-A1.
 XX
 XX 12-OCT-2001.
 PD
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 PA
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 XX WPI; 2002-043418/06.
 DR
 PT New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species -
 XX
 PS Claim 6; SEQ ID No 789; 2504pp; French.
 XX
 XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO2001/7334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.
 CC
 SQ Sequence 308 AA;
 Query Match 69.0%; Score 1048; DB 23; Length 308;
 Best Local Similarity 67.5%; Pred. No. 5.4e-88;
 Matches 208; Conservative 40; Mismatches 58; Indels 2; Gaps 1;
 QY 1 MTKTAFACQGAQYLGMRDQYDQYPTVKETIDRASQVLYGVDLRLVLTEDKLNQTRY 60
 Db 1 MTKTAFISQGAQKLGMRDQYDQYPTVKETIDRASQVLYGVDLRLVLTEDKLNQTRY 60

QY 61 TOPAILATSVAIYRLLOEKGYOPDMVAGLSGEYSALVAGSALDFEDAVAVAKRGAYME 120
 Db 61 TOPAILATSVAIYRLLOEKGYOPDMVAGLSGEYSALVAGSALDFEDAVAVAKRGAYME 120
 QY 121 EAAPADSGKWAVLNTPEVIEEACQKASEL--GVVTPANTPAQIVIAGEVAVDRAV 178
 Db 121 EAAPADSGKWAVLNTPEVIEEACQKASEL--GVVTPANTPAQIVIAGEVAVDRAV 180
 QY 179 ELIQAAGAKRLIPLKVSQPHHTALLPASPQKLAETLAQVFSDFTCPLVGNTEAAVMQKE 238
 Db 181 ELIQAAGAKRLIPLKVSQPHHTALLPASPQKLAETLAQVFSDFTCPLVGNTEAAVMQKE 240
 QY 239 DIAQLLTROVKEPVRVYESIGVMQEAQISNFIETIGPGKVLGSGFKKIDRTAHLAHVEDOA 298
 Db 241 EVKGLLTROVKEPVRVYESIGVMQEAQISNFIETIGPGKVLGSGFKKIDRTAHLAHVEDOA 300
 QY 299 SLVALLEK 306
 Db 301 SPEALINQ 308

RESULT 6
 ABP28010
 ID ABP28010 standard; Protein; 314 AA.
 AC ABP28010;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 5196.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; Gene therapy.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 XX (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelein H;
 XX
 WPI; 2002-352536/38.
 DR N-PSDB; ABB68641.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 XX Claim 1; Page 3863; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABB6044-ABN1526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to

CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 XX Sequence 314 AA;
 SQ
 Query Match 68.2%; Score 1036; DB 23; Length 314;
 Best Local Similarity 68.3%; Pred. No. 7e-87;
 Matches 209; Conservative 39; Mismatches 58; Indels 0; Gaps 0;
 QY 1 MTKTFLPAGQAQVLGMRDGYDQYPIVKETIDRASQVLGDLRYLIDTEEDKLNQTRY 60
 Db 3 MTKTFLPAGQAQVLGMRDGYDQYPIVKETIDRASQVLGDLRYLIDTEEDKLNQTRY 62
 QY 61 TOPAILATSVAIYRL-LQEKGYQPDVMAGLSLGEYSALVAGSALDFEDAVALVAKRGAYME 120
 Db 63 TOPAILATSVAIYRL-LQEKGYQPDVMAGLSLGEYSALVAGSALDFEDAVALVAKRGAYME 122
 QY 121 EAAPADSGKAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVTAGEVAVVADRAVEL 180
 Db 123 EAAPADSGKAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVTAGEVAVVADRAVEL 182
 QY 181 LQEKAGKRLIPLKVSQPPHTALLPASPQKLAETLAQVSFSDFTCPVLGNTAEAAVQKEDI 240
 Db 183 LKRGVKKLIPLVNSQPPHTALLPASPQKLAETLAQVSFSDFTCPVLGNTAEAAVQKEDI 242
 QY 241 AQLLTROVKEPVRYVESIGVWQEGAGISNFIIEGPKVLSGVFKKIDQTAHLAHVEDQASL 300
 Db 243 PELLARQVMEPVRYVESIGVWQEGAGISNFIIEGPKVLSGVFKKIDQTAHLAHVEDQASL 302
 QY 301 VALLEK 306
 Db 303 RLFLDR 308
 RESULT 7
 ABP28009
 ID ABP28009 standard; Protein; 308 AA.
 XX
 AC ABP28009;
 XX
 XX 02-JUL-2002 (first entry)
 XX
 XX Streptococcus polypeptide SEQ ID NO 5194.
 XX
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 XX Streptococcus agalactiae.
 OS
 OS WO200234771-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB04789.
 XX
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN68640.

XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 XX Claim 1; Page 3863; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (51), given in
 CC the specification. The proteins have antibacterial and antinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 XX Sequence 308 AA;
 SQ
 Query Match 65.6%; Score 996.5; DB 23; Length 308;
 Best Local Similarity 66.3%; Pred. No. 2.9e-83;
 Matches 203; Conservative 39; Mismatches 63; Indels 1; Gaps 1;
 QY 1 MTKTFLPAGQAQVLGMRDGYDQYPIVKETIDRASQVLGDLRYLIDTEEDKLNQTRY 60
 Db 1 MNKVSFLPAGQAQVLGMRDGYDQYPIVKETIDRASQVLGDLRYLIDTEEDKLNQTRY 60
 QY 61 TOPAILATSVAIYRL-LQEKGYQPDVMAGLSLGEYSALVAGSALDFEDAVALVAKRGAYM 119
 Db 61 TOPAILATSVAIYRL-LQEKGYQPDVMAGLSLGEYSALVAGSALDFEDAVALVAKRGAYM 120
 QY 120 EAAPADSGKAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVTAGEVAVVADRAVE 179
 Db 121 EAAPADSGKAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVTAGEVAVVADRAVE 180
 QY 180 LQEKAGKRLIPLKVSQPPHTALLPASPQKLAETLAQVSFSDFTCPVLGNTAEAAVQKEDI 239
 Db 181 ELKQGVKKLIPLVNSQPPHTALLPASPQKLAETLAQVSFSDFTCPVLGNTAEAAVQKEDI 240
 QY 240 AQLLTROVKEPVRYVESIGVWQEGAGISNFIIEGPKVLSGVFKKIDQTAHLAHVEDQAS 299
 Db 241 IKSLLARQVMEPVRYVESIGVWQEGAGISNFIIEGPKVLSGVFKKIDQTAHLAHVEDQAS 300
 QY 300 LVALLE 305
 Db 301 FNNLKE 306
 RESULT 8
 AAU35327
 ID AAU35327 standard; Protein; 313 AA.
 XX
 AC AAU35327;
 XX
 XX 14-FEB-2002 (first entry)
 DT
 XX Enterococcus faecalis cellular proliferation protein #614.
 DE
 XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 XX Enterococcus faecalis.
 OS
 OS WO200170955-A2.
 PN
 XX

ID XX AAY11297 standard; Protein; 137 AA.
 AC AAY11297;
 XX DT 20-MAY-1999 (first entry)
 XX DE S. pneumoniae malonyl Coa-acyl carrier protein transacylase.
 XX KW Streptococcus pneumoniae strain 0100993; vaccine; immune response;
 XX KW streptococcal infection; pneumococcal.
 XX OS Streptococcus pneumoniae.
 XX PN WO9737026-A1.
 XX PD 09-OCT-1997.
 XX PF 01-APR-1997; 97WO-US05306.
 XX PR 22-AUG-1996; 96US-0025788.
 XX PR 02-APR-1996; 96US-0014690.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
 XX PI Stodola RK;
 XX PR WPI; 1997-503111/46.
 XX DR N-PSDB; AAX30880.
 XX Nucleic acids encoding pneumococcal polypeptide(s) - useful in
 PT vaccines, drug screening, etc
 XX Claim 6; Page 305; 354pp; English.
 XX AAX30724 to AAX30946 represent genomic DNA sequences isolated from
 CC Streptococcus pneumoniae strain 0100993. These genomic DNA sequences
 CC encode the novel proteins given in AAY1114 to AAY11367. The proteins,
 CC isolated from Streptococcus pneumoniae, can be used in vaccines against
 CC streptococcal infections and in assays for identifying compounds that
 CC inhibit or activate the activity of the proteins. The antagonists can
 CC be used to treat an individual having need to inhibit a bacterial
 CC protein. Vectors expressing the proteins can be used to induce a
 CC protective immune response in mammals.
 XX SQ Sequence 137 AA;
 Query Match 44.1%; Score 670; DB 18; Length 137;
 Best Local Similarity 97.8%; Pred. No. 9.8e-54;
 Matches 134; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MTKTAFAGGGAQYLGMRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 60
 DB 1 MTKTAFAGGGAQYLRMRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 60
 QY 61 TOPAILATSVAIYRLQKGYQPDVAGISLGEYSALVAGALDPEDAVLVAKRGAYNE 120
 DB 61 TOPAILATSVAIYRLQKGYHDPVAGISLGEYSALVAGALDPEDAVLVAKRGAYNE 120
 QY 121 EAAPADSGKXMAVINTP 137
 DB 121 EAAPADSGKXMAVLHTP 137
 RESULT 11
 ID AAU34532
 AC AAU34532 standard; Protein; 309 AA.
 XX AAU34532;
 XX DT 14-FEB-2002 (first entry)

DE E. coli cellular proliferation protein #113.
 XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX OS Escherichia coli.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US09180.
 XX PR 21-MAR-2000; 2000US-191078P.
 XX PR 23-MAY-2000; 2000US-208848P.
 XX PR 26-MAY-2000; 2000US-207727P.
 XX PR 23-OCT-2000; 2000US-242578P.
 XX PR 27-NOV-2000; 2000US-253625P.
 XX PR 22-DEC-2000; 2000US-257931P.
 XX PR 16-FEB-2001; 2001US-269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen XL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX PI Yamamoto RT, Xu HH;
 XX DR WPI; 2001-611495/70.
 XX DR N-PSDB; AAS52391.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX Example 3; Seq ID No 10125; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 309 AA;
 Query Match 42.4%; Score 643.5; DB 22; Length 309;
 Best Local Similarity 45.5%; Pred. No. 8.5e-51;
 Matches 141; Conservative 60; Mismatches 102; Indels 7; Gaps 4;
 QY 1 MTKTAFAGGGAQYLGMRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQ 57
 DB 1 MTQFAFVFPGGGSGTGVNLMADMAASYFIVETFAESAALGYDLNWL--TQQGPAEELNK 58
 QY 58 TRYTOPAILATSVAIYRLQKGYQ-PDMVAGISLGEYSALVAGALDPEDAVLVAKRG 116
 DB 59 TWQTOPALLTASVALYRWVQQGGKAFAMWAGHSLGEYSALVAGVDFADAVLVMERG 118
 QY 117 AYMEEAAPADSGKXMAVINTPVEVIEACQKASELGLVVTYNTPTAQIIVAGEVAVDR 176
 DB 119 KFMQEAIVPEGTGAAVAAIIGLDDASIAKACBEAAEQGVSPVNFNSPGQVVIAGKEAVER 178
 QY 177 AVELLQEAGAKRLIPLKVSQGFHTTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAVMQ 236

Db 179 AGAACAAGAKRALPLPVSFSCALMKPAADKLAVALAKITFNAPTVPVNNVDVKET 238
 QY 237 KED-IAQLTRQVKEPVRFYBSIGVMQFAGISNFTIEIGPGKVLGSGFVKKKIDQTAHLAHE 295
 Db 239 NGDAIRDALVQLYNFQWTKSVEYMAAQGVHELVYEVGPGKVLGLTKRIVDTITASALN 298
 QY 296 DOASLVALLE 305
 Db 299 EFSMAAALE 308

RESULT 12
 AAU35382
 ID AAU35382 standard; Protein; 312 AA.
 XX
 AC AAU35382;
 DT 14-FEB-2002 (first entry)
 XX
 DE Haemophilus influenzae cellular proliferation protein #23.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Haemophilus influenzae.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS53241.
 XX
 XX New polynucleotides for the identification and development of
 PT antibiotic, comprise sequences of antisense nucleic acids -
 XX
 PS Example 3; Seq ID No 10975; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 312 AA;
 Query Match 41.9%; Score 636; DB 22; Length 312;
 Best Local Similarity 46.0%; Pred. No. 4.2e-50;
 Matches 142; Conservative 54; Mismatches 109; Indels 4; Gaps 3;
 QY 1 MYKTAPLPAQCAQVLMGRDPYDQYPIVKETIDRASQVLGYDLRYLIDT-BEDKLNQTR 59
 Db 1 MKKFAMVFFGQSGTQVGLADLATEPIVIEFKQASDALGYDLWYLVQQGPAEELNKTW 60
 QY 60 YTOPAILATSVAIYRLLOEKGYQ--PDMVAGLSLGEYSALVASGALDFEDAVLVAKRGA 117
 Db 61 QTOPALLAASVAIVRWKEKFPOLKPEVWAGHSLGEYSALVCAGVLDQDAIKLVLRGK 120
 QY 118 YMEEAAPADSGRMVAVLNTPEVIEBACQKASELGVTPPANYNTPAQIVIAGEVAVDRA 177
 Db 121 LMQQAVPEGTGAMYAIIGLDNEAIIINACKQABEGEVVSAVNFNSPGQVVIAGAKAVERA 180
 QY 178 VELLOFAGAKRLIPLKVSPPFHTALLEPASOKIASTLAQVSFSDTCTPLVGNTEA-AVMQ 236
 Db 181 AALCKEAGAKRALPLAVSVPSHCALMKPAAEQLAVTLENIQINTTISVLNNVDVKAETE 240
 QY 237 KEDIAQLTRQVKEPVRFYBSIGVMQFAGISNFTIEIGPGKVLGSGFVKKKIDQTAHLAHE 296
 Db 241 GTEIRTAIVRQLYSPVRMTETVERKMAQDGLVLAEGVPGKVLNGLTKRIVGDLQAISVND 300
 QY 297 QASLVALLE 305
 Db 301 VASFNAAVEE 309

RESULT 13
 AAU60857
 ID AAU60857 standard; Protein; 122 AA.
 XX
 AC AAU60857;
 DT 24-SEP-1998 (first entry)
 XX
 DE Amino terminal region of a FabD polypeptide embodiment.
 XX
 KW FabD protein; malonyl-CoA:ACP family; diagnosis; infection; vaccine;
 KW screen.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9822133-A1.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97WO-US20992.
 XX
 PR 18-NOV-1996; 96US-0031160.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 XX Gentry DR, Lonsdale JT, Payne DJ, Pearson SC, Van Aller G;
 XX
 XX WPI; 1998-312173/27.
 XX
 XX N-PSDB; AAV37240.
 XX
 PT New isolated Streptococcus pneumoniae FabD gene - used to develop
 PT products for the diagnosis, prevention and treatment of bacterial
 PT diseases, particularly S. pneumoniae infection
 XX
 XX Claim 23; Page 7; 45pp; English.
 CC
 CC The present sequence represents the amino terminal region of a FabD
 CC embodiment of Streptococcus pneumoniae 0100993 (NCIMB 40800). The novel
 CC FabD polypeptides are related to other proteins of the malonyl-CoA:ACP
 CC family. The products can be used for the diagnosis of Streptococcus
 CC pneumoniae infections. Vectors containing the FabD DNA sequence can be

CC administered directly to a mammal to produce the FabD peptide to
 CC provoke an antibody/T-cell response in order to prevent a disease. The
 CC Peptide can be used to screen for compounds which modulate its activity.

XX Sequence 122 AA;
 Query Match 40.4%; Score 613; DB 19; Length 122;
 Best Local Similarity 100.0%; Pred. No. 1.5e-48;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFILFAGGAQVILGMRDFYQYPIVKTIDRASQVLYDLRLYIDTEEDKLNQTRY 60
 DB 1 MTKTAFILFAGGAQVILGMRDFYQYPIVKTIDRASQVLYDLRLYIDTEEDKLNQTRY 60

QY 61 TOPAILATSVIYRLLOEKGYQPDWVAGLSLGEYSALVASCALDFEDAVALKRGAYME 120
 DB 61 TOPAILATSVIYRLLOEKGYQPDWVAGLSLGEYSALVASCALDFEDAVALKRGAYME 120

QY 121 EA 122
 DB 121 EA 122

RESULT 14
 ID ABB48893 standard; Protein; 313 AA.
 AC ABB48893;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Listeria monocytogenes protein #1597.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.
 XX
 FA (INSP) INST PASTEUR.
 XX
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihi H, Dehoux P,
 PI Dusserget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Danies J, Goebel W, Kreft J, Kuhn M, Ng B, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX
 DR WPI; 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of listeria and related bacterial infections, and
 PT related polypeptides -
 XX
 PS Claim 6; SEQ ID No 1598; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes ED-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present invention is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 313 AA;
 Query Match 38.8%; Score 588.5; DB 23; Length 313;
 Best Local Similarity 43.1%; Pred. No. 9.9e-46;
 Matches 131; Conservative 58; Mismatches 110; Indels 5; Gaps 4;

QY 1 MTKTAFILFAGGAQVILGMRDFYQYPIVKTIDRASQVLYDLRLYIDTES--DKLNQT 58
 DB 1 MTKIAFPVFGGAQVILGMRDFYQYPIVKTIDRASQVLYDLRLYIDTES--DKLNQT 58

QY 59 RYTOPAILATSVIYRLLOEKGYQPDWVAGLSLGEYSALVASCALDFEDAVALKRGAY 118
 DB 60 ENAQPALVSTSVAILRALETYGVKADYVAGHSLGEYSALVAGGFLASDAIYLVKRGSEL 119

QY 119 MEERAPADSGKMAVAVLNTPEVIEEACOKASBLG-VVTPANVNTPAQIVIAEVAVDRA 177
 DB 120 MEAVPENGAGAAVAVLGVDRDLTKITTEVTKEGDAVOLANLNCQGVISGTTAGVEKA 179

QY 178 VELLOBAGAKRLIPUKVSGPHTALLESASQKLAETLAQVFSDFCTPLVGNTEA-AVMQ 236
 DB 180 GEKAKESGAKRVLPVAVSGPHTALLESASQKLAETLAQVFSDFCTPLVGNTEA-AVMQ 239

QY 237 KEDIAQLLTQVKEPVRVFSIGVMQEGAGISNFIEIGPKVLSGFVKKIDQTAHLAHVED 296
 DB 240 KSEISDKLIKQIYSPVLWEDIVEELIKNGVDTFVEIGSKVLGLAGLIKINRDVTVLSAGD 299

QY 297 QASL 300
 DB 300 AHSV 303

RESULT 15
 AAW61018
 ID AAW61018 standard; Protein; 130 AA.
 XX
 AC AAW61018;
 XX
 DT 13-OCT-1998 (first entry)
 XX
 DE Streptococcus pneumoniae encoded polypeptide.
 XX
 KW coding region; ORF; open reading frame; antibacterial;
 KW infection; prevention; meningitis.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9819689-A1.
 XX
 PD 14-MAY-1998.
 XX
 PF 27-OCT-1997; 97WO-US19226.
 XX
 PR 01-NOV-1996; 96US-0029930.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Black MT, Hodgson JS, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Reid RH, Zarfos FN;
 XX
 DR WPI; 1998-286586/25.
 DR N-PSDB; AAV37406.

```

XX  New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT  e.g. for identifying anti-bacterial(s) for treatment and prevention
PT  of meningitis
XX
XX  Claim 11; Page 112; 130pp; English.
XX
CC  The sequence is that of the polypeptide encoded by a region isolated
CC  from S. pneumoniae which shows homology to malonyl coenzyme A-acyl
CC  carrier protein transacylase. It, or agonists of it, may be useful as an
CC  antibacterial for treatment or prevention of infection, specifically
CC  caused by S.pneumoniae (particularly meningitis) but possibly also
CC  Helicobacter pylori (ulcers and gastric cancer). It may be of particular
CC  use before insertion of an in-dwelling device or any other
CC  invasive procedure. The protein, or nucleic acid encoding
CC  it, can also be used in vaccines to induce a cellular
CC  and/or humoral immune response, or to screen for other
CC  antibacterials. The DNA may also contain flanking sequences
CC  that are potential sources of control elements for bacterial
CC  gene expression. Detecting a sequence encoding the protein
CC  can be used diagnostically, e.g. to detect a mutation for
CC  serotyping or classifying infectious agents.
XX
XX  Sequence 130 AA;
SQ
Query Match 38.5%; Score 585; DB 19; Length 130;
Best Local Similarity 93.8%; Pred. No. 6e-46;
Matches 121; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 153 VTPYNNYNTPAQIVIAGEVAVDRAVELLOEAGAKGLIPLKVGSPFHTALLEPASOKLAE 212
      |||||
Db 1 VTPYNNYNTPAQIVIAGEVAVDRAVELLOEAGAKGLIPLKVGSPFHTALLEPASOKLAE 60

QY 213 TLAQVSFSDFTCPVLGVNTEAAVMOKEDIAQLLTRQVKPEPRVYESIGVMQAGISNFTFI 272
      |||||
Db 61 TLAQVSFSDFTCPVLGVNTEAAVMOKEDIAQLLTRQVKPEPRVYESIGVMQAGISNFTFI 120

QY 273 GPGKVLSGF 281
      |||
Db 121 WTGESLVRFF 129

```

Search completed: June 11, 2003, 20:02:50
Job time : 63 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 20:01:47 ; Search time 34 Seconds
(without alignments)
264.806 Million cell updates/sec

Title: US-09-308-397-2
Perfect score: 1518
Sequence: 1 MTKTAFIFAGQGAQYLGMR.....QTAHLAHVEDQASVALLK 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTJUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	547.5	36.1	308	2	US-08-789-609A-2
2	547.5	36.1	308	4	US-09-108-517-2
3	515.5	34.0	311	4	US-09-134-001C-4824
4	376.5	24.8	4150	4	US-09-428-517-2
5	374	24.6	2259	4	US-09-413-814-70
6	374	24.6	2439	3	US-09-335-409-7
7	374	24.6	2439	4	US-09-568-102-7
8	374	24.6	2439	4	US-09-567-969-7
9	374	24.6	2439	4	US-09-568-480-7
10	374	24.6	2439	4	US-09-568-486-7
11	374	24.6	2439	4	US-09-568-472-7
12	374	24.6	2439	4	US-09-567-899-7
13	374	24.6	2475	4	US-09-413-814-48
14	372.5	24.5	5215	4	US-09-105-537-2
15	367.5	24.2	1864	2	US-08-804-227C-3
16	356	23.5	7257	3	US-09-335-409-5
17	356	23.5	7257	4	US-09-568-102-5
18	356	23.5	7257	4	US-09-567-969-5
19	356	23.5	7257	4	US-09-568-480-5
20	356	23.5	7257	4	US-09-568-486-5
21	356	23.5	7257	4	US-09-568-472-5
22	356	23.5	7257	4	US-09-567-899-5
23	351	23.1	1421	3	US-09-335-409-2
24	351	23.1	1421	4	US-09-568-102-2
25	351	23.1	1421	4	US-09-567-969-2
26	351	23.1	1421	4	US-09-568-480-2
27	351	23.1	1421	4	US-09-568-486-2

28 351 23.1 1421 4 US-09-568-472-2 Sequence 2, Appli
29 351 23.1 1421 4 US-09-567-899-2 Sequence 2, Appli
30 350.5 23.1 1841 2 US-08-804-227C-6 Sequence 6, Appli
31 349.5 23.0 6095 4 US-09-144-085-2 Sequence 2, Appli
32 348.5 23.0 2756 1 US-08-375-709-11 Sequence 11, Appli
33 348.5 23.0 2756 1 US-08-752-929-11 Sequence 11, Appli
34 348.5 23.0 2756 4 US-09-090-793-7 Sequence 7, Appli
35 344.5 22.7 4928 4 US-09-036-987A-5 Sequence 5, Appli
36 344.5 22.7 4928 4 US-09-370-700-5 Sequence 5, Appli
37 339.5 22.4 5588 4 US-09-036-987A-6 Sequence 6, Appli
38 339.5 22.4 5588 4 US-09-370-700-6 Sequence 6, Appli
39 338.5 22.3 5087 4 US-09-144-085-1 Sequence 1, Appli
40 334 22.0 343 3 US-08-858-003-32 Sequence 32, Appli
41 334 22.0 343 3 US-09-078-166-32 Sequence 32, Appli
42 334 22.0 343 4 US-08-997-457-32 Sequence 32, Appli
43 326.5 21.5 1536 4 US-09-413-814-10 Sequence 10, Appli
44 326.5 21.5 3170 4 US-09-036-987A-4 Sequence 4, Appli
45 326.5 21.5 3170 4 US-09-370-700-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-789-609A-2
; Sequence 2, Application US/08789609A
; Patent No. 5827689
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; APPLICANT: Lonsdale, John
; APPLICANT: Pearson, Stewart
; APPLICANT: Payne, David
; TITLE OF INVENTION: No. 5827689e1 FabD
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,609A
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030685
; FILING DATE: 13-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 308 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-789-609A-2

Query Match 36.1%; Score 547.5; DB 2; Length 308;
Best Local Similarity 40.6%; Fred. No. 2.1e-42;
Matches 121; Conservative 54; Mismatches 116; Indels 5; Gaps 4;

US-09-428-517-2
 ; Sequence 2, Application US/09428517
 ; Patent No. 6251636
 ; GENERAL INFORMATION:
 ; APPLICANT: Betlach, Mary C.
 ; APPLICANT: Shah, Sanjay Krishnakant
 ; APPLICANT: McDaniel, Robert
 ; APPLICANT: Tang, Li
 ; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
 ; FILE REFERENCE: 30062-20029.00
 ; CURRENT APPLICATION NUMBER: US/09/428,517
 ; CURRENT FILING DATE: 1999-10-28
 ; EARLIER APPLICATION NUMBER: 60/120,254
 ; EARLIER FILING DATE: 1999-02-16
 ; EARLIER APPLICATION NUMBER: 60/106,100
 ; EARLIER FILING DATE: 1998-10-29
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 4150
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
 ; OTHER INFORMATION: Oleandolide PKS
 US-09-428-517-2

Query Match 24.8%; Score 376.5; DB 4; Length 4150;
 Best Local Similarity 33.2%; Pred. No. 4.6e-25;
 Matches 99; Conservative 52; Mismatches 122; Indels 25; Gaps 5;
 QY 5 AFLFAGGAGVLMGRDFYQYPIVKETIDRASQVLYGIDRYLIIDTEEDK-----54
 Db 563 AFLFSGGSGQPGNGKRLHQVFGFRDALDEVCAELDTHLGRLLGPEAGPPLDRVFAER 622
 QY 55 -----LNQTRYTOPAILATSVAIYRLQEKGYQPDWAGLSLGEYSALVAGSALDDEED 107
 Db 623 GTAHSSALLSETHYTOALFALETALFLLVQWGLKPDHLACHSVGIAAARAGILDLSD 682
 QY 108 AVALVAKRGYMEAPADSGKGVAVLNTPEVIEEACQKASLGELVVTTPANYNTPAQIVT 167
 Db 683 AAELVATRGALMRSL--PGGVVMSLVQAPSEVAPILLGREAHVGL---AAVNGPDVVV 737
 QY 168 AGEVVAVDRAVELLQEAQAKELIPLKVSQPPHTALPEASQKLAETLAQVSFSDFTCPVLV 227
 Db 738 SGERGHVAATEQIURDKRSRY-LRVSHAFHSPLEPVEEFAEAVAGLTFRAPITPLV 796
 QY 228 GNTAAVYMWKEDIA--QLLTQVKEPVRFYESIGVMQAGISNFIETGPKVLSGFVK 283
 Db 797 SNLTGAPVDDRTMATPAYVVRHVREAVRFGDGIRALGKLTGSGFLEVGPDGVLTMAR 854

RESULT 5
 US-09-413-814-70
 ; Sequence 70, Application US/09413814
 ; Patent No. 6225064
 ; GENERAL INFORMATION:
 ; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
 ; APPLICANT: Bristol-Myers Squibb, Co.
 ; APPLICANT: Beyer, Stefan
 ; APPLICANT: Bloecker, Helmut
 ; APPLICANT: Brandt, Petra
 ; APPLICANT: Cino, Paul M
 ; APPLICANT: Dougherty, Brian A
 ; APPLICANT: Goldberg, Steven L
 ; APPLICANT: Hofie, Gernard
 ; APPLICANT: Mueller, Joachim
 ; APPLICANT: Reichenbach, Hans
 ; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
 ; TITLE OF INVENTION: heteropolyketide compounds
 ; FILE REFERENCE: PCT/US 99/23535
 ; CURRENT APPLICATION NUMBER: US/09/413,814
 ; CURRENT FILING DATE: 1999-10-07

; EARLIER APPLICATION NUMBER: DE 198 46 493.2
 ; EARLIER FILING DATE: 1998-10-09
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 70
 ; LENGTH: 2259
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-413-814-70

Query Match 24.6%; Score 374; DB 4; Length 2259;
 Best Local Similarity 34.8%; Pred. No. 3.2e-25;
 Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;
 QY 3 KTAFLFAGGAGVLMGRDFYDQYPIVKETIDRASQVLYGIDRYLI-----DTEED 53
 Db 373 KLAFLFAGGAGVLMGRGLWEAPAFRETDFRCVTJFDRHLHQLCEVMAEPCSSRSS 432
 QY 54 KLNQTRYTOPAILATSVAIYRLQEKGYQPDWAGLSLGEYSALVAGSALDDEED 113
 Db 433 LLDQTAFTOPALFALEYAALFRSGWVEPELVAGHSLGELVAAACVAGVFSLEDAVRLV 492
 QY 114 KRGAYMEEAAPADSGKGVAVLNTPEVIEEACQKASLGELVVTTPANYNTPAQIVAGEVVA 173
 Db 493 ARGRIM-QALFA-GGAMVSIAPAEADV---AAAVAPHAALVSIAAVNGPEQVVIAGAEKF 547
 QY 174 VDRAVELLQEAQAKELIPLKVSQPPHTALPEASQKLAETLAQVSFSDFTCPVLGNTTEAA 233
 Db 548 VQOIAAFAARGA-RTKPLHYSHAFHSPMDPMLAEFRVTVESVYRFRPSIALVSNLSGK 606
 QY 234 VMQKEDIAQ--LLTROVKEPVRFYESIGVMQAGISNFIETGPKVLSGFV 282
 Db 607 PCTDEVSAPGYVVRHVREAVRFGDGIRALGKLTGSGFLEVGPKFTLLGLV 656

RESULT 6
 US-09-335-409-7
 ; Sequence 7, Application US/09335409
 ; Patent No. 6121029
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross
 ; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/335,409
 ; CURRENT FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 2439
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-335-409-7

Query Match 24.6%; Score 374; DB 3; Length 2439;
 Best Local Similarity 34.8%; Pred. No. 3.6e-25;
 Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;
 QY 3 KTAFLFAGGAGVLMGRDFYDQYPIVKETIDRASQVLYGIDRYLI-----DTEED 53
 Db 553 KLAFLFAGGAGVLMGRGLWEAPAFRETDFRCVTJFDRHLHQLCEVMAEPCSSRSS 612
 QY 54 KLNQTRYTOPAILATSVAIYRLQEKGYQPDWAGLSLGEYSALVAGSALDDEED 113
 Db 613 LLDQTAFTOPALFALEYAALFRSGWVEPELVAGHSLGELVAAACVAGVFSLEDAVRLV 672
 QY 114 KRGAYMEEAAPADSGKGVAVLNTPEVIEEACQKASLGELVVTTPANYNTPAQIVAGEVVA 173
 Db 673 ARGRIM-QALFA-GGAMVSIAPAEADV---AAAVAPHAALVSIAAVNGPEQVVIAGAEKF 727

QY 114 KRGAYMEAPADSGKQWVAVLNTPEVIEEACQKASLGVLGVTTPANYNTPAQIIVAGEVA 173
 Db 673 ARGGLM-QALPA-GGAMVSIAPADV---AAAVAPHAALVSIAAVNGPEQVVIAGAEKF 727
 QY 174 VDRAVELLQEGAGKRLIPLKVSQGFHTALLPEPASQKLAETLAQVFSDFTCPLVGNTEAA 233
 Db 728 VQCIATAAFAARGA-RTKPLHVSHPFSLMDPMLAEAPRVVTSVYRRPSIALVSNLSGK 786
 QY 234 VMQKEDIAQ-LLTQVKEPVRFYESIGVMQEGAGISNFIETGPKVLSGFV 282
 Db 787 PCTDEVSAAGYVWRHAREAVRFADGVKALHAAGAGLFVEVGPKPTLLGLV 836

RESULT 10

US-09-568-486-7
 ; Sequence 7, Application US/09568486
 ; Patent No. 6355459
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross
 ; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/568,486
 ; CURRENT FILING DATE: 2000-05-10
 ; PRIOR APPLICATION NUMBER: 09/335,409
 ; PRIOR FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 2439
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-568-486-7

Query Match 24.6%; Score 374; DB 4; Length 2439;
 Best Local Similarity 34.8%; Pred. No. 3.6e-25;
 Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;
 QY 3 KTAFLFAGCGAQYLGMRDQYDQYPIVKTETIDRASQVGLVDRLYLI-----DTEED 53
 Db 553 KLAFLFAGCGAQVPGMGRLWEAWPAFRTDRCVTLLFDRELHQLCEVMNAEPGSSRSS 612
 QY 54 KLNQTRYTOPAILATSVAIYRLLOEKGYQPDVAGLSLGEYSALVAGSGLDFAVALVA 113
 Db 613 LLDQTAFTQPALFALEYALALFRSGVPELVAHSLGELVAACVAGVFSLEDAVRLV 672
 QY 114 KRGAYMEAPADSGKQWVAVLNTPEVIEEACQKASLGVLGVTTPANYNTPAQIIVAGEVA 173
 Db 673 ARGGLM-QALPA-GGAMVSIAPADV---AAAVAPHAALVSIAAVNGPEQVVIAGAEKF 727
 QY 174 VDRAVELLQEGAGKRLIPLKVSQGFHTALLPEPASQKLAETLAQVFSDFTCPLVGNTEAA 233
 Db 728 VQCIATAAFAARGA-RTKPLHVSHPFSLMDPMLAEAPRVVTSVYRRPSIALVSNLSGK 786
 QY 234 VMQKEDIAQ-LLTQVKEPVRFYESIGVMQEGAGISNFIETGPKVLSGFV 282
 Db 787 PCTDEVSAAGYVWRHAREAVRFADGVKALHAAGAGLFVEVGPKPTLLGLV 836

RESULT 11

US-09-568-472-7
 ; Sequence 7, Application US/09568472
 ; Patent No. 6358719
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross

; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/568,472
 ; CURRENT FILING DATE: 2000-05-10
 ; PRIOR APPLICATION NUMBER: 09/335,409
 ; PRIOR FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 2439
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-568-472-7

Query Match 24.6%; Score 374; DB 4; Length 2439;
 Best Local Similarity 34.8%; Pred. No. 3.6e-25;
 Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;
 QY 3 KTAFLFAGCGAQYLGMRDQYDQYPIVKTETIDRASQVGLVDRLYLI-----DTEED 53
 Db 553 KLAFLFAGCGAQVPGMGRLWEAWPAFRTDRCVTLLFDRELHQLCEVMNAEPGSSRSS 612
 QY 54 KLNQTRYTOPAILATSVAIYRLLOEKGYQPDVAGLSLGEYSALVAGSGLDFAVALVA 113
 Db 613 LLDQTAFTQPALFALEYALALFRSGVPELVAHSLGELVAACVAGVFSLEDAVRLV 672
 QY 114 KRGAYMEAPADSGKQWVAVLNTPEVIEEACQKASLGVLGVTTPANYNTPAQIIVAGEVA 173
 Db 673 ARGGLM-QALPA-GGAMVSIAPADV---AAAVAPHAALVSIAAVNGPEQVVIAGAEKF 727
 QY 174 VDRAVELLQEGAGKRLIPLKVSQGFHTALLPEPASQKLAETLAQVFSDFTCPLVGNTEAA 233
 Db 728 VQCIATAAFAARGA-RTKPLHVSHPFSLMDPMLAEAPRVVTSVYRRPSIALVSNLSGK 786
 QY 234 VMQKEDIAQ-LLTQVKEPVRFYESIGVMQEGAGISNFIETGPKVLSGFV 282
 Db 787 PCTDEVSAAGYVWRHAREAVRFADGVKALHAAGAGLFVEVGPKPTLLGLV 836

RESULT 12

US-09-567-899-7
 ; Sequence 7, Application US/09567899
 ; Patent No. 6383787
 ; GENERAL INFORMATION:
 ; APPLICANT: Schupp, Thomas
 ; APPLICANT: Ligon, James
 ; APPLICANT: Molnar, Istvan
 ; APPLICANT: Zirkle, Ross
 ; APPLICANT: Cyr, Devon
 ; APPLICANT: Goerlach, Joern
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
 ; FILE REFERENCE: 4-30582A
 ; CURRENT APPLICATION NUMBER: US/09/567,899
 ; CURRENT FILING DATE: 2000-05-10
 ; PRIOR APPLICATION NUMBER: 09/335,409
 ; PRIOR FILING DATE: 1999-06-17
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 2439
 ; TYPE: PRT
 ; ORGANISM: Sorangium cellulosum
 US-09-567-899-7

Query Match 24.6%; Score 374; DB 4; Length 2439;
 Best Local Similarity 34.8%; Pred. No. 3.6e-25;
 Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;
 QY 3 KTAFLFAGCGAQYLGMRDQYDQYPIVKTETIDRASQVGLVDRLYLI-----DTEED 53
 Db 553 KLAFLFAGCGAQVPGMGRLWEAWPAFRTDRCVTLLFDRELHQLCEVMNAEPGSSRSS 612

QY 54 KINQTRYQTPAILATSAIYLLQKQYQPKWAGLSLGEYSALVAGALDFEDAVLVA 113
Db 613 LLEQTAFTQPALFALEVALAALFRSGVPELVAAGHSLGELVAACVAGVFSLEDAVLV 672
QY 114 KRGAYMEAAADSKWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVIAGEVVA 173
Db 673 AGRILX-OLALFA-GGAWVSIAAPADV---AAVAPHAALVSIANVNGPEOVVIAGAEKF 727
QY 174 VDRAVELLOEAGAKRLIPKLVSGPHTALPEASQKLAETLAQVFSDFCTCLVGNTEAA 233
Db 728 VQIIAAAFARGA-RTKELHVSHPHSPMLPEAFRRVTSVTYRPSIALVSNLSGK 786
QY 234 VMQKEDIAQ-LLTRQKPEVRFYESIGVMQAGISNFTIIGPKVLGSFV 282
Db 787 PCTDEVSAPGWVHRHAREAVRFAQGVKALHAAGLFFVEVGPKEFTLLGLV 836

RESULT 13
US-09-413-814-48
; Sequence 48, Application US/094113814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloecker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 2475
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-48

Query Match 24.6%; Score 374; DB 4; Length 2475;
Best Local Similarity 33.7%; Pred. No. 3.7e-25;
Matches 99; Conservative 54; Mismatches 119; Indels 22; Gaps 8;

QY 3 KTAFLFAGGAGQVLMGRDFYDQYDPS---IVKETIDRASQVLGYDLRLYL--DTEEDKLN 56
Db 1496 RCVFLFPGGAGYPSWARDLVNCPDFALHLDPCLDQLAELLPEDPICILFGDGPADRLD 1555
QY 57 QTRYTPAILATSAIYLLQKQYQPKWAGLSLGEYSALVAGALDFEDAVLVAKRG 116
Db 1556 QYATQPLIFSVSALARWLGDFIRPDAMIGHLSIGEYVAACLAGLSLDALLVSRG 1615
QY 117 AYMEAAADSKWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVIAGEVAVDR 176
Db 1616 RUMGAA---RAMLAV-PLPEWELEERLELLAD-DRISIAVNVTASCVIAGPSEAIR 1670
QY 177 AVELLQEAGAKRLIPKLVSGPHTALPEASQKLAETLAQVFSDFCTCLVGNTE- 231
Db 1671 CAQRWAAQGL-TCTELRTSHAFSHAMPEFVPEFPHVLARVTFAPPRARWISNLDGKPID 1729
QY 232 -AAVMQKEDIAQLLTROVKEPVRFYESIGVMQAGISNFTIIGPKVLGSFVK 284
Db 1730 SAAMVQPD-----YWRHLNQPVRFEHGLSLAETHAWEVGPGRTLSSPVR 1779

RESULT 14
US-09-105-537-2
; Sequence 2, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-105-537-2

Query Match 24.5%; Score 372.5; DB 4; Length 5215;
Best Local Similarity 33.1%; Pred. No. 1.5e-24;
Matches 104; Conservative 51; Mismatches 132; Indels 27; Gaps 8;

QY 3 KTAFLFAGGAGQVLMGRDFYDQYDQYIVKETIDRASQVLGYDLRLYLID-----TEEDKL 55
Db 533 RLAVLFSGGAGQRTGGMELVAAHPAFATAFD---AAVAELDPLDRPLAELVAAGDTL 588
QY 56 NQTRYTPAILATSAIYLLQKQYQPKWAGLSLGEYSALVAGALDFEDAVLVAKR 115
Db 589 DRTVHTQPALFAVEVALHRLVSVGWTPDLLAGHSGVGEISAAHVAGVLSLRDAARLVAAR 648
QY 116 GAYMEAAADSKWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVIAGEVAVD 175
Db 649 GRIMQ--ALPEGAMVAVEASEEVLPHLAGRERLSL---AAYNGPRAVVLAGRAVL 703
QY 176 RAVELLQEAGAKRLIPKLVSGPHTALPEASQKLAETLAQVFSDFCTCLVGNTEAAVM 235
Db 704 DVAELLREQG-RTTKRLSVSHAFHSPMBPMLDFFRVVVEELDFQEPVDDVYVSTVTGLPV 762
QY 236 ---QKEDIAQLLTROVKEPVRFYESIGVMQAGISNFTIIGPKVLGSFVKIDQTAHLA 292
Db 763 TAGQWTD-PEYVWDQVRPRVFLDAVRTLUESGADTFLELPGDGVCSA-----MAADSV 815
QY 293 HVEDQASLVALLEK 306
Db 816 RQERATAVSALRK 829

RESULT 15
US-08-804-227C-3
; Sequence 3, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kunstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS


```

; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1864 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-804-227C-3

```

```

Query Match      24.2%; Score 367.5; DB 2; Length 1864;
Best Local Similarity 33.3%; Pred. No. 9.5e-25;
Matches 98; Conservative 55; Mismatches 120; Indels 21; Gaps 8;

QY      3 KTAFLFAGCAQVLGMGRDFYQYPIVKETIDRASQVIGYDL-RYLLID-----TEED 53
Db      566 RTAFLFSGGSGRAGSGRGLYRRHPVFARALDEVCAALEPHLRPLRLDLMFAEPGSPAE 625

QY      54 KLNQTRYTOPAILATSVATYRLLOEKGYQDVMVAGLSLGEYSALVAGSALDFEDAVALVA 113
Db      626 PLDRTEFTQPALFALQTALFRLAEHGLRAELCGHSVGEIAAAHAGVLTLPDAARLVA 685

QY      114 KRGAYMEEAPADSGKXAVLNTPEVIIEACQKASELGVWTFPANNYNTPAQIVIAGEVVA 173
Db      686 ARGRLM-QALPA-GGAMAALRATAETIAPLLERRAGELAL---AAVNGPSSVWVSGDEAA 740

QY      174 VDRAVELLQE--AGAKRLIPLKVGSPFHTALLEPASQKLAETLAQVRSDFTCPLVGNTE 231
Db      741 V---LELEQWRAEGREARRLAVSHAFHSFRMDGMLTQFDRVARTLTFPPTIPLVSLT 797

QY      232 AAVNQKEDI--AQLLTRQVKPEPVRYESIGVMQEGAGISNFIIGPKVLSGFVK 283
Db      798 GTPVTETCTADHWVRQAREPVRFELDAMRTLEADGIDTFVELGPDGVLSAMAR 851

```

Search completed: June 11, 2003, 20:06:35
Job time : 35 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 20:04:42 ; Search time 309 Seconds
(without alignments)
102.238 Million cell updates/sec

Title: US-09-308-397-2

Perfect score: 1518

Sequence: 1 MTKTAFLPAGQAQYLMGR.....QTAHLARVEDQASVALLLEK 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1514	99.7	306	10	US-09-815-242-13580
2	937	61.7	313	10	Sequence 13580, A
3	643.5	42.4	309	10	Sequence 10920, A
4	636	41.9	312	10	Sequence 10125, A
5	556.5	36.7	302	10	Sequence 10975, A
6	556.5	36.7	311	10	Sequence 5460, A
7	556.5	36.7	312	10	Sequence 12130, A
8	416.5	27.4	309	10	Sequence 11894, A
9	414.5	27.3	1402	9	Sequence 11290, A
10	376.5	24.8	4150	9	Sequence 166, App
11	374	24.6	2439	9	Sequence 2, Appli
12	372.5	24.5	5215	9	Sequence 7, Appli
13	372.5	24.5	5215	9	Sequence 2, Appli
14	372.5	24.5	5215	9	Sequence 2, Appli
15	372.5	24.5	5215	10	Sequence 2, Appli
16	371	24.4	1616	9	Sequence 262, App
17	356	23.5	7257	9	Sequence 5, Appli
18	351	23.1	1421	9	Sequence 2, Appli
19	346.5	23.0	2756	9	Sequence 7, Appli

20	348	22.9	395	9	US-10-074-045-46	Sequence 46, Appl
21	340.5	22.4	426	9	US-10-124-800-10	Sequence 10, Appl
22	340.5	22.4	2910	9	US-10-124-800-2	Sequence 2, Appli
23	334	22.0	343	9	US-09-735-556-32	Sequence 32, Appl
24	324	21.3	4613	9	US-09-860-846-31	Sequence 31, Appl
25	324	21.3	4613	9	US-09-988-384B-31	Sequence 31, Appl
26	324	21.3	4613	9	US-09-836-821-31	Sequence 31, Appl
27	324	21.3	4613	10	US-09-861-289-31	Sequence 31, Appl
28	324	21.3	11877	9	US-09-860-846-6	Sequence 6, Appli
29	324	21.3	11877	9	US-09-836-821-6	Sequence 6, Appli
30	324	21.3	11877	10	US-09-861-289-6	Sequence 6, Appli
31	324	21.3	12199	9	US-09-988-384B-6	Sequence 31, Appl
32	306	20.2	328	9	US-09-735-556-31	Sequence 4, Appli
33	275.5	18.1	1832	9	US-10-014-717-4	Sequence 6, Appli
34	274	18.1	3798	9	US-10-014-717-6	Sequence 46, Appl
35	254	16.7	2563	9	US-09-836-705-46	Sequence 4, Appli
36	253.5	16.7	3519	9	US-09-808-880-4	Sequence 3, Appli
37	246	16.2	3816	9	US-09-808-880-3	Sequence 261, App
38	235.5	15.5	1827	9	US-09-712-363-261	Sequence 34, Appl
39	235	15.5	345	9	US-09-735-556-34	Sequence 6666, Ap
40	221	14.6	1610	9	US-09-738-626-6666	Sequence 35, Appl
41	217.5	14.3	1562	9	US-09-860-846-35	Sequence 35, Appl
42	217.5	14.3	1562	9	US-09-988-384B-35	Sequence 35, Appl
43	217.5	14.3	1562	9	US-09-836-821-35	Sequence 35, Appl
44	217.5	14.3	1562	10	US-09-861-289-35	Sequence 35, Appl
45	212	14.0	1346	9	US-09-860-846-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-13580
; Sequence 13580, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13580
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13580

Query Match 99.7%; Score 1514; DB 10; Length 306;
Best Local Similarity 99.7%; Pred. No. 5.2e-124;

Matches 305; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 60
 Db 1 MTKTAFAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 60

QY 61 TOPAILATSVAYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKGAYME 120
 Db 61 TOPAILATSVAYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKGAYME 120

QY 121 EAAPADSGMWAVLNTPVEVIEEACQKASELGVVTPANYNTPAQIVIAGEVAVDRAVEL 180
 Db 121 EAAPADSGMWAVLNTPVEVIEEACQKASELGVVTPANYNTPAQIVIAGEVAVDRAVEL 180

QY 181 LOEAGAKELIPLKVSQPHHTALLEPASQKLAETTLAQVSFDFTCPLVGNTEAAVMQKEDI 240
 Db 181 LOEAGAKELIPLKVSQPHHTALLEPASQKLAETTLAQVSFDFTCPLVGNTEAAVMQKEDI 240

QY 241 AQLLTQVKEPVRVYESIGVMOEAGISNFIETGPKVLSGFVKKIDQTAHLAHVEDQASL 300
 Db 241 AQLLTQVKEPVRVYESIGVMOEAGISNFIETGPKVLSGFVKKIDQTAHLAHVEDQASL 300

QY 301 VALLEK 306
 Db 301 VALLEK 306

RESULT 2

US-09-815-242-10920
 ; Sequence 10920, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011a
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10920
 ; LENGTH: 313
 ; TYPE: PRF
 ; ORGANISM: Enterococcus faecalis
 US-09-815-242-10920

Query Match 61.7%; Score 937; DB 10; Length 313;
 Best Local Similarity 61.9%; Pred. No. 9,3e-74;
 Matches 187; Conservative 40; Mismatches 75; Indels 0; Gaps 0;

QY 3 KTAFLPAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 62
 Db 3 KTAFLPAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 62

Db 6 KTAFLPAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQTRY 65

QY 63 PAILATSVAYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKGAYMEEA 122
 Db 66 PAILTUSVAFYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKGAYMTEA 125

QY 123 APADSGMWAVLNTPVEVIEEACQKASELGVVTPANYNTPAQIVIAGEVAVDRAVELIQ 182
 Db 126 APOGTGKMWAVMNAEVEIEKACQASAFIVAPANYNTPAQIVIAGEVAVDRAVELIQ 185

QY 183 EAGAKELIPLKVSQPHHTALLEPASQKLAETTLAQVSFDFTCPLVGNTEAAVMQKEDI 242
 Db 186 EAGKEMIFLNVSGPHHTALLOPASKKLAQDLAKLNFQTMQIPVSNITABIMPQEAICA 245

QY 243 LLTQVKEPVRVYESIGVMOEAGISNFIETGPKVLSGFVKKIDQTAHLAHVEDQASLVA 302
 Db 246 LLEKQVMSAVRPFEDSIETKAMVNGVTMIEVGPKLTGTGFKKIDKTIEMHVRVEDVATLTE 305

QY 303 LL 304
 Db 306 TL 307

RESULT 3

US-09-815-242-10125
 ; Sequence 10125, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011a
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10125
 ; LENGTH: 309
 ; TYPE: PRF
 ; ORGANISM: Escherichia coli
 US-09-815-242-10125

Query Match 42.4%; Score 643.5; DB 10; Length 309;
 Best Local Similarity 45.5%; Pred. No. 3.3e-48;
 Matches 141; Conservative 60; Mismatches 102; Indels 7; Gaps 4;

QY 1 MTKTAFAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQ 57
 Db 1 MTKTAFAGQAQYGLMGGRDFDYQYPIVKETIDRASQVLGYDLRYLIDTEEDKLNQ 58

QY 58 TMTQPAIATSVAYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKG 116
 Db 58 TMTQPAIATSVAYRLLOEKGYQPDVWAGLSLGEYSALVASCALDPEDAVAVAKG 116

Db 59 TWQTQALLTASVALYRWQOQGGKAPAMMAGHSLGYSALVCAGVIDFADAVRLVEMRG 118
 QY 117 AYMEBAAPAGQGMVAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDR 176
 Db 119 KFMQEAPEGTGMAAIIIGLDASIAKACEBAAGQVSPVNFSPQGVVIAHKEAVER 178
 QY 177 AVELLOEAGAKRLPLKVSFPHFHTALLEPASQKLAETLAQVSPSDFCTCLVGNTEAAVMQ 236
 Db 179 AGAACKAAGAKRALPLVSPVSHCALMKPAADKLAVELAKITFNAPTVPVNNVNDVKCET 238
 QY 237 KED-IAQLLTRQVKPEVRFYESIGVMQEBAGISNFIETGPGKVLSGFVKKIDQTAHLAHVE 295
 Db 239 NGDAIRDALVQLYNPQWTKSVYMAAQGVHELYEVPKGKVLTKRLTKRIVDTLTASALN 298
 QY 296 DOASLVALLE 305
 Db 299 EPSAAVALE 308

RESULT 4

US-09-815-242-10975
 ; Sequence 10975, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10975
 ; LENGTH: 312
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-815-242-10975

Query Match 41.9%; Score 636; DB 10; Length 312;
 Best Local Similarity 46.0%; Pred. No. 1.5e-47;
 Matches 142; Conservative 54; Mismatches 109; Indels 4; Gaps 3;
 QY 1 MTKTAFIFAGGAQVLMGRDFYQYPIVKTIDRASQVLYDRLYIDT-EEDKLNQTR 59
 Db 1 MKKTFAMVPGGSGTQVGMADLATEYPIVETPKQASDALGYDLYWLVQGPAAELNKTW 60
 QY 60 YTOPAILATSVAIYRLQEKGYQ--PDMVAGLSLGEYSALVASCALDFEDAVALKRGA 117
 Db 61 QTOPALLAASVAIYRWKEKFPQKPPEVMAGHSLGEYSALVCAGVLPDQAIKLVLRGK 120
 QY 118 YMEBAAPADSGKMAVNLNTPVEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDR 177

Db 121 LMQQAVPEGTGMVAIIGLDNEAIINACKQABEGEVSAVNFNSPGQVVIAGAKAVERA 180
 QY 178 VELLOEAGAKRLPLKVSFPHFHTALLEPASQKLAETLAQVSPSDFCTCLVGNTEA-AVMQ 236
 Db 181 AALCKEAGAKRALPLAVSPVSHCALMKPAABQLAVTLENIQINTPTISLVANNVDVKAETE 240
 QY 237 KEDIAQLLTRQVKPEVRFYESIGVMQEBAGISNFIETGPGKVLSGFVKKIDQTAHLAHVED 296
 Db 241 GTEITALTALVRLYSPVRMTETVETKVAQDGLVLAIEVGGKVLNGLTKRLIVGDLQALSVND 300
 QY 297 QASLVALLE 305
 Db 301 VASFNAAVEE 309

RESULT 5

US-09-815-242-5460
 ; Sequence 5460, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE OF INVENTION: Prokaryotes
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5460
 ; LENGTH: 302
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-09-815-242-5460

Query Match 36.7%; Score 556.5; DB 10; Length 302;
 Best Local Similarity 40.9%; Pred. No. 1.2e-40;
 Matches 122; Conservative 54; Mismatches 117; Indels 5; Gaps 4;
 QY 1 MTKTAFIFAGGAQVLMGRDFYQYPIVKTIDRASQVLYDRLYIDT-EEDKLNQTR 59
 Db 1 MSKTALIFPGGAQVLMGRDFYQYPIVKTIDRASQVLYDRLYIDT-EEDKLNQTR 60
 QY 60 YTOPAILATSVAIYRLQEKGYQ--PDMVAGLSLGEYSALVASCALDFEDAVALKRGA 119
 Db 61 NTQPALHTSSALLAAL--KNLNPFTMGHSLGEYSALVAAVLSFEBAVKIVRKGQIM 118
 QY 120 EBAAPADSGKMAVNLNTPVEVIEEACQK-ASELGVVTPANTYNTPAQIVIAGEVAVDR 178
 Db 119 AQAPFTGVGMAVILGLDFDKVDEICKSLSSDDKIEEPANINCPQIVVSGHKALIDELV 178
 QY 179 ELLQEAGAKRLPLKVSFPHFHTALLEPASQKLAETLAQVSPSDFCTCLVGNTEA-AVMQ 237

```
Db 179 EKGSLGAKRWPLAVSGFFHSSLMKVIEEDFSSYINQFEHDAKFPVQVNVNAQGETDK 238
QY 238 EDIAQLLRQVKEPVRFYES:GVMQEGAGISNFIIGPGKVLGSGFVKKIDQTAHLAHVE 295
Db 239 EVIKSNMVKQLYSPVQFINSTEWLIDQGVDFHFIIGPGKVLGSLGKIKNRDVKLTSIQ 296

RESULT 6
US-09-815-242-12130
; Sequence 12130, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12130
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12130

Query Match 36.7%; Score 556.5; DB 10; Length 311;
Best Local Similarity 40.9%; Pred. No. 1.2e-40;
Matches 122; Conservative 54; Mismatches 117; Indels 5; Gaps 4;

QY 1 MTKTAPFAGGQAQYLGMRDIFYDQYPIVKETIDRASQVLGYDLRYLIDTEED-KLNQTR 59
Db 4 MSKTAIFPGGQAQYGVMAQQLYNNNDQATEILTSAAINTLDFDILEMTFTDDEGKLGETE 63

QY 60 YTOPAILATSVAIYELLOEKGYQPDVAGLSLGEYSALVSGALDFEDAVALKVAKRGAYM 119
Db 64 NTQPAALLTHSSALLAAL--KMLNPDFTMGHSLGETSSLSVAADVUSFSDAVKIVRKGQLM 121

QY 120 EEAAPADSGKMWAVLNTPEVEIEACQK-ASELGWVTPYANTNTPAQIVTAGEVAVDRAV 178
Db 122 AQAPFTGVGMAAVLGLDFDKVDEICKSLSDDKIIEPANINCPGQIVVSHKALIDELV 181

QY 179 ELLQEGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSFSDFTCTPLVGNTEA-AVMQK 237
Db 182 EKGSLGAKRWPLAVSGFFHSSLMKVIEEDFSSYINQFEHDAKFPVQVNVNAQGETDK 241

QY 238 EDIAQLLRQVKEPVRFYES:GVMQEGAGISNFIIGPGKVLGSGFVKKIDQTAHLAHVE 295
Db 242 EVIKSNMVKQLYSPVQFINSTEWLIDQGVDFHFIIGPGKVLGSLGKIKNRDVKLTSIQ 299

RESULT 7
US-09-815-242-11894
; Sequence 11894, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11894
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11894

Query Match 36.7%; Score 556.5; DB 10; Length 312;
Best Local Similarity 44.5%; Pred. No. 1.2e-40;
Matches 126; Conservative 51; Mismatches 103; Indels 3; Gaps 3;

QY 5 AFLFAGGQAQYLGMRDIFYDQYPIVKETIDRASQVLGYDLRYLIDT-BEDKLNQTRYTOP 63
Db 6 AFVFPQGSQSLGMLAEALGAQQAQLVRDTFAEASEALGYDLWALVQNGPEERLNQTDKTOP 65

QY 64 AILATSVAIYRL-LOEKGYPDMVAGLSLGEYSALVSGALDFEDAVALKVAKRGAYMEEA 122
Db 66 AILTSIALWRLWLAEAGNAPAFVAGHSLGEYSALVVAESLAFADAVKLVERGQLMQQA 125

QY 123 APADSGKMWAVLNTPEVEIEACQKASELGWVTPYANTNTPAQIVTAGEVAVDRAVELLQ 182
Db 126 VPAGQGGMAAILGLEDDADVLAAEAAQOEVSVAVNFNAPGQVVIAGAAAVERAIEACK 185

QY 183 EAGAKELIPLKVSQGFHTALLEPASQKLAETLAQVSFSDFTCTPLVGNTEAAVMQKED-IA 241
Db 186 ARGAKRAVALPVSVSHCELMPAAEQFAASVESLQWQAPKISLVQNVSAAPDLDILR 245

QY 242 QLLTRQVKEPVRFYESIGVMQEGAGISNFIIGPGKVLGSGFVKK 284
Db 246 RLLLAQLYSPVAVWVESIQLLAEKGVTELVECGPGKVLAGLNRR 288

RESULT 8
US-09-815-242-11290
; Sequence 11290, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
```

```

; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 166
; LENGTH: 1402
; TYPE: PR1
; ORGANISM: Mycobacterium tuberculosis
; US-09-712-363-166

Query Match      27.3%; Score 414.5; DB 9; Length 1402;
Best Local Similarity 39.9%; Pred. No. 2.3e-27;
Matches 115; Conservative 39; Mismatches 117; Indels 17; Gaps 8

Qy 3 KTAFLPAGGQAQYLGMRDFFYDQVPVKTETIDRASQVLGYDLY-LIDT-----EEDKLNQ 57
Db 630 KTAFLVAGGQWLGMGSELYAYPVFAEALDAVDDELDRHLYRPLRDYVINGHDQDLNT 689
Qy 58 TRYTOPALATSVAIYLLQEKGYOPDMVAGLSLGEVSALVASGALDPEDAVALVAKRGA 117
Db 690 TFEAQLFAFNEVALYKLLWSGVRFGLVGHVGLAAHVAAGALCLPDAAMLVAARGR 749
Qy 118 YMEEAAPADSGKWAVLNTPVVEIEACQKASELG-VVTPANYNTPAQIVIAGEVVAVDR 176
Db 750 LM-QALPA-GGAMFAVQAREDEV-----APMLGHVDVSTAANGVSPASVVISGAHDVSA 800
Qy 177 AVELLQEAQAKRLPLKVGSPFHTALLEPASQKLAETLAQVSPSDFTCPLVGN-TEAAVM 235
Db 801 IADRLRGQG-RVVRHLAVSHAFSALMEPMIABFTVAARELSVGLPTIPVINSVTGQLVA 859
Qy 236 QKEDIAQLLTQVKEPVRVFESIGVMQEAQISNPIGIGPKVLGSFVK 283
Db 860 DPFASADYWARHIRAVRFGDSVSAHCACASRFFIEVFGGGGLTSLIE 907

RESULT 10
US-09-808-880-2
; Sequence 2, Application US/09808880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shan, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

```

```
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: Oleandolide PKS
US-09-808-880-2

Query Match      24.8%; Score 376.5; DB 9; Length 4150;
Best Local Similarity 33.2%; Pred. No. 2.1e-23;
Matches 99; Conservative 52; Mismatches 122; Indels 25; Gaps 5;

QY 5 AFLEAGGAGYLGMRGRDYDQYPIVKETIDRASQVIGYDLRYLIDTDEDK-----54
Db 563 AFLEGGGQQGPMGRHLHQVFFGRDALDEVCAELDTLGRLLGPPAGPPLRVMFAER 622
QY 55 -----LNQTRYTOPALATSVAIYRLLQKGYQPMVAGLSLGEYSALVSGALDFED 107
Db 623 GTASHALLSETHYTOAALFALETAFLRLVQWGLKPDHLAGHSVGEIAAAHAAGILDSD 682
QY 108 AVALYAKGAYMEEAAPADSGKMWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVI 167
Db 683 AAELVATRGALMSL--PGGVMLSVQAPSEVAPLLLGREAHVGL---AAVNGPDPAVVV 737
QY 168 AGEVAVDRAVELLOEAGAKRLIPLKVSPPHTALLESASQKLAETLAQVSFSDFTCPVLV 227
Db 738 SCERGHVAAIEQLLDRGRKSRV-LRVSHAFHSPLMPEVLEFAEAVAGLTFRAPTTPLV 796
QY 228 GNTAAVNMQKEDIA--QLLTQVKEPVRFYESIGVMQAGISNFIETGPGKVLGSGFVK 283
Db 797 SNLTGAPVDDRTMATPAYVWRHREAVRFGDGRALGKLTGSGFLEVPDGVLTAMAR 854

RESULT 11
US-10-014-717-7
; Sequence 7, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Rees
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2439
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-10-014-717-7

Query Match      24.6%; Score 374; DB 9; Length 2439;
Best Local Similarity 34.8%; Pred. No. 1.7e-23;
Matches 101; Conservative 38; Mismatches 135; Indels 16; Gaps 6;

QY 3 KTAFLPAGGAGYLGMRGRDYDQYPIVKETIDRASQVIGYDLRYLI-----DTEED 53
Db 553 KTAFLPAGGAGYLGMRGRDLMEAWFAFRETFTDRCVTLLFDRELHPLCEVWMAEPGSSRSS 612
QY 54 KLNQTRYTOPALATSVAIYRLLQKGYQPMVAGLSLGEYSALVSGALDFEDAVLVA 113
Db 613 LLDQTAFTQPALFALEYALALFRSGVGEPELVAGHSIGELVAAACVAGVPSLEDAVLV 672
QY 114 KEGAYMEEAAPADSGKMWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVAGEVVA 173
Db 673 ARGRLM-QALPA-GGAMYSIAAPEADV---AAAVAPHAALYSIAAVNGPEQVVIAGAEKF 727
```

```
QY 174 VDRAVELLOEAGAKRLIPLKVSPPHTALLESASQKLAETLAQVSFSDFTCPVLGNTAA 233
Db 728 VQOIAAFAARGA-RTKPLHVSFAHSPMLDPMLEAFRRVTSVYRRFSIALVSNLSGK 786
QY 234 VMQKEDIAQ-LLTROVKEPVRFYESIGVMQAGISNFIETGPGKVLGSGFV 282
Db 787 PCTDEVSAFGYVWRHREAVRFGDGVKALHAAGAGLFVEVGPKPTLLGLV 836

RESULT 12
US-09-860-846-2
; Sequence 2, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
US-09-860-846-2

Query Match      24.5%; Score 372.5; DB 9; Length 5215;
Best Local Similarity 33.1%; Pred. No. 6.4e-23;
Matches 104; Conservative 51; Mismatches 132; Indels 27; Gaps 8;

QY 3 KTAFLPAGGAGYLGMRGRDYDQYPIVKETIDRASQVIGYDLRYLI-----TEEDKL 55
Db 533 RLAVLPSGGGQAGRTGTGMELVAAHPAFATFD---AVAAELDPLDRPLAELVAGDTL 588
QY 56 NQTRYTOPALATSVAIYRLLQKGYQPMVAGLSLGEYSALVSGALDFEDAVLVAKR 115
Db 589 DRTVHTQPALFAVEVALHRLVESWGTTPDLLAGHSVGEISAAHVAGVLSLRDAARLVAR 648
QY 116 GAYMEEAAPADSGKMWAVLNTPEVIEEACQKASELGVVTPANYNTPAQIVAGEVVAVD 175
Db 649 GRLMQ--ALPEGGAMVAVEASEEVLPLAGRELSL---AAVNGPRAVLAGAERAVL 703
QY 176 RAVELLOEAGAKRLIPLKVSPPHTALLESASQKLAETLAQVSFSDFTCPVLGNTAAVM 235
Db 704 DVAELLREQG-RTKRLSVSHAFHSPMLDPMLEAFRRVVEELDFOEPRVDVVSTVTGLPV 762
QY 236 ---QKEDIAQLLTROVKEPVRFYESIGVMQAGISNFIETGPGKVLGSGFVKIDQTAHLA 292
Db 763 TAGQWTD-PEYVWDQVRREPVRFLDAVRTLEESGADTFLELGPDGVCSA-----MAADSV 815
QY 293 HVEDQASLVALLEK 306
Db 816 RDQERATAVSALRK 829

RESULT 13
US-09-988-384B-2
; Sequence 2, Application US/09988384B
; Publication No. US20030073824A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.5360US1
```

BEST AVAILABLE COPY

```

; CURRENT APPLICATION NUMBER: US/09/988,384B
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: PCT/US99/14398
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 53
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-988-384B-2

Query Match      24.5%; Score 372.5; DB 9; Length 5215;
Best Local Similarity 33.1%; Pred. No. 6.4e-23;
Matches 104; Conservative 51; Mismatches 132; Indels 27; Gaps 8;

QY 3 KTAFLFAGGGAQYLGMRDFFDYQYPIVKETIDRASQVGLGYDLRYLID-----TEEDKL 55
DB 533 RLAVLFSGGGAQRTGMGMELVAAHPAFATFD----AAVAELDPLLDRPLAELVAAGDTL 588

QY 56 NOTRYTOPAILATSVAIYRLLOEKGYOPDMVAGLSLGEYSALVAGSALDGEDAVALVAKR 115
DB 589 DRTVHTQPALFAVEALHRLVSVSGVTPDLLAGHSVGEISAAHVAGVLSLRDAARLVAAR 648

QY 116 GAYMEEAAPADSGKMWAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVIAGEVAVD 175
DB 649 GRIMQ--ALPEGGAAMVAEASEEVLPHLAGRERLSL---AAVNGPRAVVLAGAERAVL 703

QY 176 RAVELLOEAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVPSDFTCPLVGNTEAAVM 235
DB 704 DVAELLREQG--RRTKRLSVSHAFSPMLPEMLDDFRVVEELDFQEPRVDVSTVTGLPV 762

QY 236 ---QKEDIQALLTROVKEPVRVYESIGVMOBAGISNFIEIGPGKVLGSGFVKKIDQTAHLA 292
DB 763 TAGQWTD--PEYWDQVRRPVRFLDAVRTLEESGADTFLELGPDCGYCSA-----MAADSV 815

QY 293 HVEDQASLVALLEK 306
DB 816 RDOEAATAVSALRK 829

RESULT 14
US-09-836-821-2
; Sequence 2, Application US/09936821
; Publication No. US20030087405A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/836,821
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 2001-04-17
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-836-821-2

Query Match      24.5%; Score 372.5; DB 9; Length 5215;
Best Local Similarity 33.1%; Pred. No. 6.4e-23;
Matches 104; Conservative 51; Mismatches 132; Indels 27; Gaps 8;

QY 3 KTAFLFAGGGAQYLGMRDFFDYQYPIVKETIDRASQVGLGYDLRYLID-----TEEDKL 55
DB 533 RLAVLFSGGGAQRTGMGMELVAAHPAFATFD----AAVAELDPLLDRPLAELVAAGDTL 588

QY 56 NOTRYTOPAILATSVAIYRLLOEKGYOPDMVAGLSLGEYSALVAGSALDGEDAVALVAKR 115
DB 589 DRTVHTQPALFAVEALHRLVSVSGVTPDLLAGHSVGEISAAHVAGVLSLRDAARLVAAR 648

QY 116 GAYMEEAAPADSGKMWAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVIAGEVAVD 175
DB 649 GRIMQ--ALPEGGAAMVAEASEEVLPHLAGRERLSL---AAVNGPRAVVLAGAERAVL 703

QY 176 RAVELLOEAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVPSDFTCPLVGNTEAAVM 235
DB 704 DVAELLREQG--RRTKRLSVSHAFSPMLPEMLDDFRVVEELDFQEPRVDVSTVTGLPV 762

QY 236 ---QKEDIQALLTROVKEPVRVYESIGVMOBAGISNFIEIGPGKVLGSGFVKKIDQTAHLA 292
DB 763 TAGQWTD--PEYWDQVRRPVRFLDAVRTLEESGADTFLELGPDCGYCSA-----MAADSV 815

QY 293 HVEDQASLVALLEK 306
DB 816 RDOEAATAVSALRK 829

```

```

QY 56 NOTRYTOPAILATSVAIYRLLOEKGYOPDMVAGLSLGEYSALVAGSALDGEDAVALVAKR 115
DB 589 DRTVHTQPALFAVEALHRLVSVSGVTPDLLAGHSVGEISAAHVAGVLSLRDAARLVAAR 648

QY 116 GAYMEEAAPADSGKMWAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVIAGEVAVD 175
DB 649 GRIMQ--ALPEGGAAMVAEASEEVLPHLAGRERLSL---AAVNGPRAVVLAGAERAVL 703

QY 176 RAVELLOEAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVPSDFTCPLVGNTEAAVM 235
DB 704 DVAELLREQG--RRTKRLSVSHAFSPMLPEMLDDFRVVEELDFQEPRVDVSTVTGLPV 762

QY 236 ---QKEDIQALLTROVKEPVRVYESIGVMOBAGISNFIEIGPGKVLGSGFVKKIDQTAHLA 292
DB 763 TAGQWTD--PEYWDQVRRPVRFLDAVRTLEESGADTFLELGPDCGYCSA-----MAADSV 815

QY 293 HVEDQASLVALLEK 306
DB 816 RDOEAATAVSALRK 829

RESULT 15
US-09-861-289-2
; Sequence 2, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 5215
; TYPE: PRT
; ORGANISM: Streptomyces venezuelae
; US-09-861-289-2

Query Match      24.5%; Score 372.5; DB 10; Length 5215;
Best Local Similarity 33.1%; Pred. No. 6.4e-23;
Matches 104; Conservative 51; Mismatches 132; Indels 27; Gaps 8;

QY 3 KTAFLFAGGGAQYLGMRDFFDYQYPIVKETIDRASQVGLGYDLRYLID-----TEEDKL 55
DB 533 RLAVLFSGGGAQRTGMGMELVAAHPAFATFD----AAVAELDPLLDRPLAELVAAGDTL 588

QY 56 NOTRYTOPAILATSVAIYRLLOEKGYOPDMVAGLSLGEYSALVAGSALDGEDAVALVAKR 115
DB 589 DRTVHTQPALFAVEALHRLVSVSGVTPDLLAGHSVGEISAAHVAGVLSLRDAARLVAAR 648

QY 116 GAYMEEAAPADSGKMWAVLNTPEVEIEACOKASELGVVTPANYNTPAQIVIAGEVAVD 175
DB 649 GRIMQ--ALPEGGAAMVAEASEEVLPHLAGRERLSL---AAVNGPRAVVLAGAERAVL 703

QY 176 RAVELLOEAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVPSDFTCPLVGNTEAAVM 235
DB 704 DVAELLREQG--RRTKRLSVSHAFSPMLPEMLDDFRVVEELDFQEPRVDVSTVTGLPV 762

QY 236 ---QKEDIQALLTROVKEPVRVYESIGVMOBAGISNFIEIGPGKVLGSGFVKKIDQTAHLA 292
DB 763 TAGQWTD--PEYWDQVRRPVRFLDAVRTLEESGADTFLELGPDCGYCSA-----MAADSV 815

QY 293 HVEDQASLVALLEK 306
DB 816 RDOEAATAVSALRK 829

```


Search completed: June 11, 2003, 20:17:37
Job time : 310 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 19:47:12 ; Search time 44 Seconds
(without alignments)
668,571 Million cell updates/sec

Title: US-09-308-397-2
Perfect score: 1518
Sequence: 1 MTKTAFLEAGGAGYLGMR.....QTAHLAHVEDQASLVALLEK 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pirl:.*
2: pirl2:.*
3: pirl3:.*
4: pirl4:.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1518	100.0	306	2 F95048	malonyl CoA-acyl c
2	1515	99.8	306	2 D97919	[acyl-carrier-prot
3	1048	69.0	308	2 E86721	hypothetical prote
4	713.5	47.0	308	2 G97338	malonyl CoA-acyl c
5	669.5	44.1	317	2 H69620	[acyl-carrier-prot
6	657.5	43.3	312	2 G82128	[acyl-carrier-prot
7	654	43.1	313	2 D81961	malonyl CoA-acyl c
8	643.5	42.4	309	1 E44856	[acyl-carrier-prot
9	640.5	42.2	309	2 F90812	malonyl-CoA-[acyl-
10	640.5	42.2	309	2 B85672	malonyl-CoA-[acyl-
11	636	41.9	312	2 E64051	[acyl-carrier-prot
12	628.5	41.4	309	2 A10194	[acyl-carrier-prot
13	612.5	40.3	309	2 AC0642	malonyl CoA-acyl c
14	611.5	40.3	307	2 T44432	[acyl-carrier-prot
15	594.5	39.2	326	2 F70376	[acyl-carrier-prot
16	588.5	38.8	313	2 AH1300	malonyl CoA-acyl c
17	586	38.6	313	2 AH1672	malonyl CoA-acyl c
18	570	37.5	308	2 A81972	[acyl-carrier-prot
19	566	37.3	308	2 B81026	[acyl-carrier-prot
20	556.5	36.7	312	2 T12019	[acyl-carrier-prot
21	551.5	36.3	308	2 A89896	malonyl CoA-acyl c
22	533.5	35.1	317	2 E82776	[acyl-carrier-prot
23	531	35.0	308	2 A86528	malonyl acyl carri
24	529	34.8	305	2 A75334	[acyl-carrier-prot
25	528	34.8	308	2 E72096	[acyl-carrier-prot
26	526	34.7	293	2 G72334	[acyl-carrier-prot
27	492	32.4	331	2 C97493	malonyl-CoA-acyl c
28	492	32.4	331	2 AE2711	malonyl-CoA:acyl c
29	485	31.9	314	2 AH3436	[acyl-carrier-prot

30	479	31.6	292	2 AH1836	malonyl coenzyme A
31	479	31.6	314	2 H87456	malonyl CoA-acyl c
32	468	30.8	314	2 D97839	hypothetical prote
33	464.5	30.6	308	2 F81895	[acyl-carrier-prot
34	461.5	30.4	306	2 D81428	[acyl-carrier-prot
35	459	30.2	293	2 S75355	probable [acyl-car
36	457.5	30.1	400	2 T44805	malonyl-CoA transa
37	451.5	29.7	308	2 G71538	[acyl-carrier-prot
38	445	29.3	319	2 D71533	[acyl-carrier-prot
39	439.5	29.0	367	2 T00580	probable [acyl-car
40	426	28.1	650	2 C69678	involved in polyke
41	425	28.0	288	2 A69678	probable polyketid
42	420.5	27.7	309	2 F71976	probable [acyl-ket
43	419	27.6	3413	2 T17467	xifamycin polyketi
44	416.5	27.4	309	2 B64531	probable [acyl-car
45	414.5	27.3	1402	2 D70634	probable polyketid

ALIGNMENTS

RESULT 1
F95048
malonyl CoA-acyl carrier protein transacylase [imported] - Streptococcus pneumoniae (st
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: F95048
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
neon, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrisor
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Cross-references: GB:AE005672; PIDN:AAK74583.1; PID:G14971890; GSPDB:GN00164; TIGR:SE
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95048
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <R>
A:Cross-references: GB:AE005672; PIDN:AAK74583.1; PID:G14971890; GSPDB:GN00164; TIGR:SE
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0420
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-me

Query Match	100.0%;	Score	1518;	DB	2;	Length	306;
Best Local Similarity	100.0%;	Pred. No.	3.3e-99;	Indels	0;	Gaps	0;
Matches	306;	Conservative	0;	Mismatches	0;		
Qy	1	MTKTAFLEAGGAGYLGMRDFYDQYPIVKETIDRASQVLGYDLRLYLIDTEEDKLNQTRY	60				
Db	1	MTKTAFLEAGGAGYLGMRDFYDQYPIVKETIDRASQVLGYDLRLYLIDTEEDKLNQTRY	60				
Qy	61	TQPAILTATVAIYRLLOEKGYQPDWAGLSLGEYSALVASGALDFEDAVALKRGAYME	120				
Db	61	TQPAILTATVAIYRLLOEKGYQPDWAGLSLGEYSALVASGALDFEDAVALKRGAYME	120				
Qy	121	EAAPADSGKMAVLTNPVEVIEBACOKASELGVVTPANYNTPAQIVIAGEVAVDRAVEL	180				
Db	121	EAAPADSGKMAVLTNPVEVIEBACOKASELGVVTPANYNTPAQIVIAGEVAVDRAVEL	180				
Qy	181	LOEAGAKRLIPLKVGSPFHTALLEPASOKLAETLAQVSPDSFTCPPLVGNTEAAVMQKEDI	240				
Db	181	LOEAGAKRLIPLKVGSPFHTALLEPASOKLAETLAQVSPDSFTCPPLVGNTEAAVMQKEDI	240				
Qy	241	AQLTTRQKVPVRPVEYISIGVMQAGISNFIETGPGKVLGVFKKIDQTAHLAHVEDQASL	300				
Db	241	AQLTTRQKVPVRPVEYISIGVMQAGISNFIETGPGKVLGVFKKIDQTAHLAHVEDQASL	300				
Qy	301	VALLEK 306					
Db	301	VALLEK 306					

RESULT 2
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) [imported] - Streptococcus pne
D97919
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: D97919
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain Re.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: D97919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99184.1; PID:gl5457941; GSPDB:GNC00174
C:Genetics:
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal
C:Keywords: acyltransferase; coenzyme A

Query Match 99.8%; Score 1515; DB 2; Length 306;
Best Local Similarity 99.7%; Pred. No. 5.4e-99;
Matches 305; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTRY 60
DB 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTRY 60
QY 61 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAYME 120
DB 61 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAYME 120
QY 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 180
DB 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 180
QY 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 180
DB 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 180
QY 181 LQAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQKEDI 240
DB 181 LQAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQKEDI 240
QY 241 AQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 300
DB 241 AQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 300
QY 301 VALLEK 306
DB 301 VALLEK 306

RESULT 3
E86721
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: E86721
R:Bolesin, A.; Winkler, P.; Mauger, S.; Jalllon, O.; Malarme, K.; Weissbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: E86721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <STO>
A:Cross-references: GB:AE0051176; PID:gl2723692; PIDN:AAK04871.1; GSPDB:GNC0146
A:Experimental source: strain IL1403
C:Genetics:
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal
Query Match 69.0%; Score 1048; DB 2; Length 308;

Best Local Similarity 67.5%; Pred. No. 3.1e-66;
Matches 208; Conservative 40; Mismatches 58; Indels 2; Gaps 1;
QY 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTRY 60
DB 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTRY 60
QY 61 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAYME 120
DB 61 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAYME 120
QY 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 178
DB 121 EAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 178
QY 179 ELLOEAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQK 238
DB 179 ELLOEAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQK 238
QY 239 DIAQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 298
DB 239 DIAQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 298
QY 299 SLVALLEK 306
DB 301 SFEALINQ 308
RESULT 4
G97338
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: G97338
R:Noelling, J.; Brston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le
Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G97338
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK81498.1; PID:gl5026671; GSPDB:GNC0168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
C:Gene: CAC3575
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-m

Query Match 47.0%; Score 713.5; DB 2; Length 308;
Best Local Similarity 48.5%; Pred. No. 8.8e-45;
Matches 147; Conservative 56; Mismatches 97; Indels 3; Gaps 3;
QY 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTR 59
DB 1 MTKTAFAGGAGQYLGMRDFYDQYPIVKETIDRASQVLGYDLRLIDTEEDKLNQTR 59
QY 60 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAY 118
DB 60 TQPAIATSVAIYRLLEKGYQPDWVAGSLGEYSALVAGSALDFFEDAVALKRGAY 118
QY 119 MEAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 178
DB 119 MEAPADSGKQVAVLNTPEVIEEACQKASEL--GWTPNANTPAQIVIAGEVAVDRAV 178
QY 179 ELLOEAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQK 237
DB 179 ELLOEAGAKRLIPLKVSQVGFHTALLESQKLAETLAQVSDFTCPVLGVNTEAAVMQK 237
QY 238 EDIAQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 297
DB 238 EDIAQLLTROVKPEVPFYESIGWQVAGISNFIETIGPGKVLGSGVFKKIDOTAHLAHVDDQ 297
QY 241 DOIKGLLKKQVMSVSRWEDTIRRMDDGVDTFIELGPGKTLSSFIKKINRMTIFNIEKA 300
DB 241 DOIKGLLKKQVMSVSRWEDTIRRMDDGVDTFIELGPGKTLSSFIKKINRMTIFNIEKA 300

QY 298 ASL 300
Db 301 BDL 303

RESULT 5
H69620
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) fabd [validated] - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: H69620; T46632
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Enian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
A.; Jech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpe, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A59580; MUID:98044033; PMID:9384377
A:Accession: H69620
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-317 <KUN>
A:Cross-references: GB:Z99112; GB:AL009126; NID:G2633902; PIDN:CAB13463.1; PID:G2633962
A:Experimental source: strain 168
R:Cronan, J.E.; Morbidoni, H.R.; de Mendoza, D.
J. Bacteriol. 178, 4794-4800, 1996
A:Title: Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosynt
A:Reference number: Z23107; MUID:96326321; PMID:8759840
A:Accession: T46632
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-123, 'GCRLEKQW', 132-317 <CRO>
A:Cross-references: EMBL:U59433; NID:G1502418; PIDN:AAC4306.1; PID:G1502420
A:Experimental source: strain 168
C:Genetics:
A:Gene: fabD
A:Map position: 135-145 degrees
C:Function:
A:Description: EC 2.3.1.39 [validated, MUID:96326321]
A:Pathway: lipid synthesis
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal
C:Keywords: acyltransferase; coenzyme A; fatty acid biosynthesis
F:5-289/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F:91/Active site: Ser (covalent substrate-binding) #status predicted
F:201/Active site: His #status predicted

Query Match 44.1%; Score 669.5; DB 2; Length 317;
Best Local Similarity 47.2%; Pred. No. 1:1e-39;
Matches 143; Conservative 47; Mismatches 110; Indels 3; Gaps 3;

QY 1 MTKTAPLAGAQAQVLMGRDQYDQYPIVKETIDRASOVLGVDLRYLI-DTEEDKLNQTR 59
Db 1 MSKIAFLPFGQSQIGVKGELVEQVPAKRLFDKDEADLETLSLIFEQDAELITY 60

QY 60 YTPQAILATSAIYRLLEKQVPMVAGLSLGEYSALVAGSALDGFEDAVALVAKRAYM 119
Db 61 NAAQPAALLTSTAIVLEKFKESGITPDTAGHSLGEYSALVAGSALDGFEDAVTVVRKGEFM 120

QY 120 BEAAPADSGKMAVNTPEVIEEACQKASELG-VTPANVTYPAQIVTAGEVAVDRAV 178
Db 121 NEAVPAGEGAMAIIGMDAEALUKVTDKVTBEGNLVQANLNCQGIIVISTAKGVELAS 180

QY 179 ELLOEAGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVFSDFTCPLVGNTEAAVM-QK 237

Db 181 ELAKENGAKRAITPLEVSGPFHSELMPAAEKLEVLDAACDIKDADVPVISNVASDVMTKE 240
QY 238 EDIAQLTRQVKEPVRPFYESTIGWMOEAGISNFIETGPKVLSGFGVKKIDQTAHLAHVEDQ 297
Db 241 ADIKEKLEQLSPVRFESINKLIAEGVTTFIEIGPKVLSGLVKVNRRLKTTIAVSDP 300
QY 298 ASL 300
Db 301 EII 303

RESULT 6
G82128
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) [similarity] - Vibrio cholerae
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-Aug-2001
C:Accession: G82128
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.
Chardson, D.; Emlaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
I. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82128
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <HEI>
A:Cross-references: GB:AE004276; GB:AE003852; NID:G9656555; PIDN:AAF95170.1; GSPDB:GN00
A:Experimental source: serogroup O1; strain N16951; biotype El Tor
C:Genetics:
A:Gene: VC2022
A:Map position: 1
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-ma
C:Keywords: acyltransferase; coenzyme A
F:97/Active site: Ser (covalent substrate-binding) #status predicted
F:206/Active site: His #status predicted

Query Match 43.3%; Score 657.5; DB 2; Length 312;
Best Local Similarity 47.4%; Pred. No. 7.5e-39;
Matches 146; Conservative 58; Mismatches 97; Indels 7; Gaps 4;

QY 1 MTKTAPLAGAQAQVLMGRDQYDQYPIVKETIDRASOVLGVDLRYLI-DTEEDKLNQTR 59
Db 6 MSKIAFLPFGQSQIGVKGELVEQVPAKRLFDKDEADLETLSLIFEQDAELITY 65

QY 60 YTPQAILATSAIYRLLEKQY-QPMVAGLSLGEYSALVAGSALDGFEDAVALVAKRAY 118
Db 66 RIQPALLAASVAIRWVWQQLGLEQPAVLAGHSLGEYSALVAGSALDGFEDAVLRLGQL 125

QY 119 MBEAPADSGKMAVNTPEVIEEACQKASELG-VTPANVTYPAQIVTAGEVAVDRAV 178
Db 126 MQQAVPAGTGAMVYAIIGLEDEIAKACADAQGEVVPVNFNSPGQVVIAGKDAVERAG 185

QY 179 ELLOEAGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVFSDFTCPLVGNTEAAVMQKE 238
Db 186 VLCKEAGAKRLPLPVSVFHCALMKPADELAKTLAELEFNAPQIPVINVVD--VVAET 243

QY 239 D---IAQLTRQVKEPVRPFYESTIGWMOEAGISNFIETGPKVLSGFGVKKIDQTAHLAHE 295
Db 244 DPKIKDALIRQLSPVRFESINKLIAEGVTTFIEIGPKVLSGLVKVNRRLKTTIAVSDP 303

QY 296 DQASLVAL 303
Db 304 DVASLDVAV 311

RESULT 7
D83961
malonyl CoA-acyl carrier protein transacylase fabd [imported] - Bacillus halodurans (st:
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83961
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir

Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: D83961
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-313 <SPO>
A/Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA806211.1; GSPDB:GN00
A/Experimental source: strain C-125
C/Genetics:
A/Gene: fabD
C/Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-malonyltransferase
Query Match 43.1%; Score 654; DB 2; Length 313;
Best Local Similarity 45.6%; Pred. No. 1.3e-38;
Matches 141; Conservative 59; Mismatches 106; Indels 4; Gaps 4;
QY 1 MTKTAEFLPAGQAGYQGLMGSRDFDQYPIVKETIDRASQVGLYDL-RLYLDTEDKLNQTR 59
DB 1 MAKVAFLEPQGGSQSGVMGSELLSEK-AKEIFDEADERLGYSLSSIMFEGPEKLRTE 59
QY 60 YTOPAILATSVAIYRLLOEKGYQDMVAGLSLGEYSALVAGSALDPEDADVALVAKGAYM 119
DB 60 NTOPALLTSTAVLSVIREYIGRPDYTAGHSLGEYSALVAGSLTFADAVYVHRGLFM 119
QY 120 EEAPADSGRWAVLNTPEVIEEACQKASELG-VVTPANYNTPAQIVIAGEVWADRAV 178
DB 120 EEAVPREGAMAILGWERDELQVTKRVTEAGAVVELANLNCPOIIVSGSAEGVEQAS 179
QY 179 ELLOEAGAKRLIPLKVSPPHTALLEPASOKLAETLAQVSDFTCPVLGNTTEAAVMQK- 237
DB 180 EEAKEAGAKRVIPLQVSGPFFHSLMKPAEKLAEVLADIAADAPPVIANVTADLVQKA 239
QY 238 EDIAQLLTROVKEPVPFVYFISGVQWQAGISNFTIEIGPGKVLGSFVKKIDQTAHLAHVEDQ 297
DB 240 ADIRSLIEGVYGFVWEDTVRMLGLVDTFVEIGSNVLGSLVRKVRNVFVSVD 299
QY 298 ASLVALLEK 306
DB 300 ASIEAMVKK 308
RESULT 8
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) fabD [validated] - Escherichia
C/Species: Escherichia coli
C/Date: 04-Mar-1993 #sequence revision 13-Jan-1995 #text_change 01-Mar-2002
C/Accession: B41856; A42147; S20443; A64853
R/Verwoert, I.I.; Varbre, E.C.; van der Linden, K.H.; Nijkamp, H.J.; Stuitje, A.R.
J. Bacteriol. 174, 2851-2857, 1992
A>Title: Cloning, nucleotide sequence, and expression of the Escherichia coli fabD gene,
A/Reference number: A41856; MUID:92234941; PMID:1314802
A/Accession: B41856
A/Molecule type: DNA
A/Residues: 1-309 <VER>
A/Cross-references: GB:M87040; NID:g145985; PIDN:AAA23742.1; PID:g145887
A/Note: Sequence extracted from NCBI backbone (NCBI:97135; NCBI:97148)
R/Rawlings, M.; Cronan Jr., J.E.
J. Biol. Chem. 267, 5751-5754, 1992
A>Title: The gene encoding Escherichia coli acyl carrier protein lies within a cluster
A/Reference number: A42147; MUID:92210530; PMID:1556094
A/Accession: A42147
A/Molecule type: DNA
A/Residues: 289-309 <RAW>
A/Cross-references: GB:M484991; NID:g145879; PIDN:AAA23738.1; PID:g145880
R/Magnuson, K.; Oh, W.; Larson, T.J.; Cronan Jr., J.E.
FEBS Lett. 299, 262-266, 1992
A>Title: Cloning and nucleotide sequence of the fabD gene encoding malonyl coenzyme A-
A/Reference number: S20443; MUID:92183950; PMID:1339356
A/Accession: S20443
A/Molecule type: DNA
A/Residues: 1-309 <WAG>
A/Cross-references: EMBL:211565; NID:g41363; PIDN:CAA77658.1; PID:g41364

A/Experimental source: strain K-12
A/Note: amino end of the mature protein confirmed by protein sequencing
R/Ruch, F.E.; Vagelos, P.R.
J. Biol. Chem. 248, 8095-8106, 1973
A>Title: Characterization of a malonyl-enzyme intermediate and identification of the m-
A/Reference number: A55383; MUID:74030718; PMID:4584823
A/Contents: annotation; active site
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: A64853
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-309 <BIAT>
A/Cross-references: GB:AE00210; GB:U00096; NID:g1787332; PIDN:AAC74176.1; PID:g1787334
A/Experimental source: strain K-12, substrain MG1655.
C/Genetics:
A/Gene: fabD; tfpA
A/Map position: 24 min
C/function:
A/Description: catalyzes the reaction of malonyl-CoA with acyl carrier protein to form
A/Pathway: fatty acid biosynthesis
C/Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-m-
C/Keywords: acyltransferase; coenzyme A; fatty acid biosynthesis
F:2-309/Product: [acyl-carrier-protein] S-malonyltransferase #status experimental <WAT>
F:5-289/Domain: [acyl-carrier-protein] S-malonyltransferase homology <WAT>
F:92/Active site: Ser (covalent substrate-binding) #status experimental
F:201/Active site: His #status predicted
Query Match 42.4%; Score 643.5; DB 1; Length 309;
Best Local Similarity 45.5%; Pred. No. 7.1e-38;
Matches 141; Conservative 60; Mismatches 102; Indels 7; Gaps 4;
QY 1 MTKTAEFLPAGQAGYQGLMGSRDFDQYPIVKETIDRASQVGLYDLRLVLDTEE---DKLNQ 57
DB 1 MTPQAFVFPQGGSQGTGMLADMAASYPIVEETFAEASALGYSALGDMAL--TQGPFAELNK 58
QY 58 TRYTOPAILATSVAIYRLLOEKGYQ-PDMVAGLSLGEYSALVAGSALDPEDADVALVAKRG 116
DB 59 TWQTOPALLTSTAVLSVIREYVWQGGKAPAMWAGHSLGEYSALVAGVIDFADAVRLVEMRG 118
QY 117 AYMEAPADSGRWAVLNTPEVIEEACQKASELG-VVTPANYNTPAQIVIAGEVWADRAV 176
DB 119 KFMQEAPEVPGTGMAALIGLDDASIAKACERAAEGQVSPVNFSPQGVVIAGHKEAVER 178
QY 177 AVELLOEAGAKRLIPLKVSPPHTALLEPASOKLAETLAQVSDFTCPVLGNTTEAAVMQ 236
DB 179 AGAACKAAGAKRVLPLQVSGPFFHSLMKPAEKLAEVLADIAADAPPVIANVTADLVQKA 238
QY 237 KED-IAQLLTROVKEPVPFVYFISGVQWQAGISNFTIEIGPGKVLGSFVKKIDQTAHLAHVE 295
DB 239 NGDAIRDALVRLQVNPVQWTKSVYMAAQGVHLYEVGPGKVLGTLTKRIVDTLTASALN 298
QY 296 DQASLVALLE 305
DB 299 EPSAFAAALE 308
RESULT 9
malonyl-CoA-[acyl-carrier-protein] transacylase [imported] - Escherichia coli (strain C
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence revision 18-Jul-2001 #text_change 03-Aug-2001
C/Accession: F90812
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: F90812
A/Status: preliminary

AI0194
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) [imported] - Yersinia pestis
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 11-Jan-2002
C:Accession: AI0194
R:Farhili, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-raraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AI0194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <XUR>
A:Cross-references: GB:AL590842; PIDN:CAC90420.1; PID:gl5979636; GSPDB:GN00175
C:Genetics:
A:Gene: fabD
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal
C:Keywords: acyltransferase; Coenzyme A

Query Match 41.4%; Score 628.5; DB 2; Length 309;
Best Local Similarity 45.5%; Pred. No. 8.1e-37;
Matches 140; Conservative 56; Mismatches 109; Indels 3; Gaps 3;
Query 1 MTKTAFLEAGGAGQAYLGMGRDFDYQYPIVKETIDRASQVLGYDLRYLIDT-EEDKLNQTR 59
Db 1 MSKFAVFPQGSQSLMADLAQAQPIVEATFSEASSVLGYDLWQLVQGPABELNKTW 60

Query 60 YTOPAILATSVAIYRLLOEKGYO-PDMVAGLSLGEYSALVAGSALDGFEDAVLAKRAY 118
Db 61 QTOPALLTASVAIYRLLOEKGYO-PDMVAGLSLGEYSALVAGSALDGFEDAVLAKRAY 120

Query 119 MBEAAPADSGKQAVLNTPVEIEEACQKASELGVTPYNTPAQIVIAAGEVAVDRAV 178
Db 121 MQEAVPEGTGAMVAITGLDNESIAKACEEAGQVSPVNFNSPGQVVIAGNKDAVERAG 180

Query 179 ELIQEAGAKELIPLKVSQPFHTALLESASQKLAETLAQVSFSDTCTPLVGNTER-ANMOK 237
Db 181 AACKAAGAKALPLVSPVSHCALMKPAADKLAVALKIDFQEPPLFPVNVNVDKAEVSP 240

Query 238 EDIAQLTRQVKPEPVRFVYESIGVMQEAGISNFIETGPGKVLGSGFVKKIDQTAHLAHVEDQ 297
Db 241 VAIRSAVLRYLPVNPVWTSVEFIAAEGVELLEIGPGKVLGTLTKRIVDSLAANAANDV 300

Query 298 ASLVALLIE 305
Db 301 VTLSALE 308

RESULT 13
AC0642
malonyl CoA-acyl carrier protein transacylase [imported] - Salmonella enterica subsp. en
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC0642
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Mule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AC0642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08318.1; PID:gl6502364; GSPDB:GN00176
C:Genetics:
A:Gene: STY1233
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-mal

Query Match 40.3%; Score 612.5; DB 2; Length 309;
Best Local Similarity 43.9%; Pred. No. 1.1e-35;
Matches 139; Conservative 62; Mismatches 97; Indels 17; Gaps 5;
Query 1 MTKTAFLEAGGAGQAYLGMGRDFDYQYPIVKETIDRASQVLGYDLRYLIDT---DKLNQ 57
Db 1 MTQFAVFPQGSQSLMADLAQAQPIVEATFSEASSVLGYDLWQLVQGPABELNKTW 58

Query 58 TRYTOPAILATSVAIYRLLOEK-GYOPDMVAGLSLGEYSALVAGSALDGFEDAVLAKRG 116
Db 59 TWOTQALLTASVALWRVWQGGKMPALMAGHSLGEYSALVAGSALDGFEDAVLAKRG 118

Query 117 AYMEAAPADSGKQAVLNTPVEIEEACQKASELGVTPYNTPAQIVIAAGEVAVDR 176
Db 119 KFMQEAPEGTGAMVAITGLDNESIAKACEEAGQVSPVNFNSPGQVVIAGHKEAVER 178

Query 177 AVELQEAGAKELIPLKVSQPFHTALLESASQKLAETLAQVSFSDTCTPLVGN-----T 230
Db 179 AGAACKAAGAKALPLVSPVSHCALMKPTADKLAVALAKITFSAPTVPVNVNVDKCVET 238

Query 231 EAAVMQKEDIQALLTRQVKPEPVRFVYESIGVMQEAGISNFIETGPGKVLGSGFVKKIDQTAH 290
Db 239 DAAA-----IRDALVRLQVNPVQWTKSVFIAAQQVHELYEYVGPGLVGLTKRIVDTLT 293

Query 291 LAHVEDQASLVALL 304
Db 294 ASALNEPAALSAAL 307

RESULT 14
T44433
[acyl-carrier-protein] S-malonyltransferase (EC 2.3.1.39) fabD [similarity] - Moritella
C:Species: Moritella marina
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jun-2000
C:Accession: T44433
R:Morita, N.; Ueno, A.; Tanaka, M.; Ohgiya, S.; Hoshino, T.; Kawasaki, K.; Yumoto, I.;
Biotechnol. Lett. 21, 641-646, 1999
A:Title: Cloning and sequencing of clustered genes involved in fatty acid biosynthesis
A:Reference number: Z22768
A:Accession: T44433
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-307 <MOR>
A:Cross-references: EMBL:AB021978; PIDN:BAA85255.1
A:Experimental source: ATCC 15381
C:Genetics:
A:Gene: fabD
C:Superfamily: [acyl-carrier-protein] S-malonyltransferase; [acyl-carrier-protein] S-ma
C:Keywords: acyltransferase; coenzyme A
F:5-289/Domain; [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F:92/Active site: Ser (covalent substrate-binding) #status predicted
F:201/Active site: His #status predicted

Query Match 40.3%; Score 611.5; DB 2; Length 307;
Best Local Similarity 45.7%; Pred. No. 1.2e-35;
Matches 139; Conservative 59; Mismatches 101; Indels 5; Gaps 4;
Query 1 MTKTAFLEAGGAGQAYLGMGRDFDYQYPIVKETIDRASQVLGYDLRYLID--TEBDKLNQ 58
Db 1 MSKYAVFPQGSQSLMADLAQAQPIVEATFSEASSVLGYDLWQLVQGPABELNKTW 59

Query 59 RYTOPAILATSVAIYRLLOEK-GYOPDMVAGLSLGEYSALVAGSALDGFEDAVLAKRG 117
Db 60 HITOPALLTASVALWRVWQGGKMPALMAGHSLGEYSALVAGSALDGFEDAVLAKRG 119

Query 118 YMEAAPADSGKQAVLNTPVEIEEACQKASELGVTPYNTPAQIVIAAGEVAVDRA 177
Db 120 LMQQAVPOGICGAMAARVIGLNDALIAACATAAEDEVVSAVNFSFGQVVIAGNKAUVRA 179

Query 178 VELLQEAGAKELIPLKVSQPFHTALLESASQKLAETLAQVSFSDTCTPLVGNTE-AAVMQ 236
Db 180 SELCVGARRVMPPLVSPVSHCSLMKPADELKLALEKVTFTNTPVTKLINNVDAAPVD 239

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 19:45:37 ; Search time 36 Seconds
(without alignments)
352.549 Million cell updates/sec

Title: US-09-308-397-2
Perfect score: 1518
Sequence: 1 MTKTALFACGQAYLONGR.....QTAHLHVZDQSLVALLKX 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	669.5	44.1	317	1 FABD_BACSU	P71019 bacillus su
2	638.5	42.1	308	1 FABD_ECOLI	P25715 escherichia
3	636	41.9	312	1 FABD_FAEIN	P25712 haemophilus
4	611.5	40.3	308	1 FABD_SALTY	O85140 salmonella
5	459	30.2	293	1 FABD_SYNY3	P73242 synechocyst
6	305.5	20.1	3491	1 ERY1_SACER	Q03131 saccharopol
7	274.5	18.1	302	1 FABD_MYCTU	Q10501 mycobacteri
8	269	17.7	2181	1 STCA_EMENI	Q12397 emericella
9	266.5	17.6	1986	1 WKEMENI	Q03149 emericella
10	254	16.7	2109	1 KXSL_ASPPA	Q12053 aspergillus
11	253.5	16.7	3519	1 OL56_STRAT	Q07017 streptomyce
12	246	16.2	3567	1 ERY2_SACER	Q03132 saccharopol
13	242.5	16.0	3172	1 ERY3_SACER	Q03133 saccharopol
14	224	14.8	1538	1 PSB_MYCTU	Q10978 mycobacteri
15	208.5	13.7	1876	1 PSB_MYCTU	Q10977 mycobacteri
16	205	13.5	2037	1 FAS1_CANAL	P34731 c fatty aci
17	199.5	13.1	2505	1 FAS1_FAT	P12785 rattus norv
18	196.5	12.9	2076	1 FAS1_YARLI	P34229 y fatty aci
19	195.5	12.9	1774	1 MSAS_PENPA	P22367 penicillium
20	186	12.3	2110	1 MCAS_MYCBO	Q02251 mycobacteri
21	184	12.1	2051	1 FAS1_YEAST	P07149 s fatty aci
22	173.5	11.4	2073	1 FAS1_SCHPO	Q5uug0 s fatty aci
23	169	11.1	2504	1 FAS1_HUMAN	P49327 homo sapien
24	167.5	11.0	2511	1 FAS1_CHICK	P12276 gallus gall
25	149.5	9.8	1914	1 STCK_EMENI	Q00706 emericella
26	108	7.1	398	1 PKK_LACLA	Q9C1W1 lactococcus
27	101.5	6.7	563	1 CGL1_ORYSA	P71814 oryza sativ
28	99.5	6.6	431	1 ENO_CLOAB	Q97152 clostridium
29	99.5	6.6	502	1 ATPA_GUITH	O78475 guillardia
30	99.5	6.6	505	1 ATPA_CYAPA	P28080 cyanophora
31	99.5	6.6	506	1 ATPA_ANASP	P12405 anabaena sp
32	98.5	6.5	756	1 YASI_RHISN	P5652 rhizobium s
33	97.5	6.4	3396	1 POLG_DENIS	P33478 d genome po

ALIGNMENTS

RESULT 1

ID	FABD_BACSU	STANDARD;	PRT;	317 AA.
AC	P71019: O34463;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).			
GN	FABD.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=96326321; PubMed=8759840;			
RT	"Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosynthesis genes."			
RT	J. Bacteriol. 178:4794-4800(1996).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98195738; PubMed=9534248;			
RT	"A 28 kbp segment from the spoVM region of the Bacillus subtilis 168 genome."			
RT	Microbiology 144:801-805(1998).			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98044033; PubMed=9384377;			
RT	Kunst F, Ogatawara N, Moszer I, Albertini A.M., Alloni G., Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S., Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J., Conerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haleth J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningstein G., Krcgh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche R., Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconelli E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,			

```

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256(1997).
CC -!- CARBOLYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
CC malonyl-[acyl-carrier protein].
CC -!- PATHWAY: Fatty acid biosynthesis.
CC -!- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
CC TRANSACYLASE ACTIVITY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U59433; AAC44306.1; -
DR EMBL; Y13937; CAA74249.1; -
DR EMBL; Z99112; CAB13463.1; -
DR HSSP; P25715; 1MLA.
DR Subtilist; BG11836; fabD.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; fabD.
DR Pfam; PF00698; Acyl_transf; 1.
DR TIGRFAMs; TIGR00128; fabD; 1.
KW Fatty acid biosynthesis; Transferase; Complete proteome.
FT ACT SITE 91 BY SIMILARITY.
FT ACT SITE 201 BY SIMILARITY.
FT CONFLICT 124 131 VPAGEGAM -> GCRLAKEOM (IN REF. 1).
SQ SEQUENCE 317 AA; 34035 MW; 191AE828B1C91D7F CRC64;

Query Match 44.1%; Score 669.5; DB 1; Length 317;
Best Local Similarity 47.2%; Pred. No. 3.8e-39;
Matches 143; Conservative 47; Mismatches 110; Indels 3; Gaps 3;

QY 1 MTKTAFIFAGGAGVGLMGDFDYDQVPIVKETIDRASVGLGYDLVLI-DTEEDKLNQTR 59
Db 1 MSKTAFLFPGGSGFINGKELHEVQVPAARLFDDEADTLETKSLIFEGDAEETLTY 60
QY 60 YTOPAILATSVAIYRLLOEKGYQDMVAGLSLGEYSALVASGALDFEDAVLVAKRGAYM 119
Db 61 NAOQALLTTSTAVLEKFKESGITPDFTAGHSLGEYSALVAAGALSFKDAVYTVRKRGPEM 120
QY 120 EERAPADSGKMAVINTFVEVIEACOKASELG-VYTPPNYNTPAQIVTAGEVVAVDRAV 178
Db 121 NEAVPAGEGAAALIGMDAELKQVTDKVTTEGNLVQLANLCPGQIVISGTAKGVELAS 180
QY 179 ELLOEAGAKRLIPLKVSQPFHTALLEPASOKLAETLAQVFSDFTCPLVGNTEAAVM-QK 237
Db 181 ELAKENGAKRAIPLVSPGFHSELMAPAEKKEVLDAACDIDKADVPLSNVSADVWTEK 240
QY 238 EDIAQLITRQVKEPVRYEIGVQWQAGISNFIIEGKVLGSKVFKIDQTAHLAHEVQ 297
Db 241 ADIKEKLEIQLYSPVFEESINKLIAEGVTFTEIGPKVLGLVKVARRLKNTIAVSDP 300
QY 298 ASL 300
Db 301 ETI 303

RESULT 2
FABD_ECOLI STANDARD; PRT; 308 AA.
ID FABD_ECOLI
AC P25715;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).

```

```

GN FABD OR TPPA OR B1092.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=K12;
RX MEDLINE=92183950; PubMed=1339356;
RA Magnuson K., Oh W., Larson T.J., Cronan J.E. Jr.;
RT "Cloning and nucleotide sequence of the fabD gene encoding malonyl
RL coenzyme A-acyl carrier protein transacylase of Escherichia coli.";
PEBS Lett. 239:262-266(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92234941; PubMed=1314802;
RA Verwoert I.I., Verbree E.C., van der Linden K.H., Nijkamp H.J.,
RT Stuitje A.R.;
RC "Cloning, nucleotide sequence, and expression of the Escherichia coli
RT fabD gene, encoding malonyl coenzyme A-acyl carrier protein
RT transacylase."; J. Bacteriol. 174:2851-2857(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95214624; PubMed=7700236;
RA Bouquin N., Tempete M., Holland I.B., Seror S.J.;
RT "Resistance to trifluoroperazine, a calmodulin inhibitor, maps to the
RT fabD locus in Escherichia coli.";
RL Mol. Gen. Genet. 246:628-637(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden W.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [6]
RP SEQUENCE OF 288-308 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92210530; PubMed=1556094;
RA Cronan J.E. Jr., Rawlings M.;
RT "The gene encoding Escherichia coli acyl carrier protein lies within
RT a cluster of fatty acid biosynthetic genes.";
RL J. Biol. Chem. 267:5751-5754(1992).
RN [7]
RP SEQUENCE OF 1-11.
RC STRAIN=K12 / W3110;
RA Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
RA Hochstrasser D.F.;
RL Submitted (SEP-1994) to the SWISS-PROT data bank.
RN [8]
RP SEQUENCE OF 1-10.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded

```

in the genome of *Escherichia coli* K-12." ;
[9]
X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
MEDLINE=95285670; PubMed=7768883;
Serre L., Verbree E.C., Dauter Z., Smitte A.R., Derewenda Z.S.;
"The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase
at 1.5-A resolution. Crystal structure of a fatty acid synthase
component";
J. Biol. Chem. 270:12961-12964(1995).
CC -!- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
CC malonyl-[acyl-carrier protein].
CC -!- PATHWAY: Fatty acid biosynthesis.
CC -!- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
CC TRANSACYLASE ACTIVITY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M87040; AAA23742.1; -
CC EMBL; Z11565; CAA77658.1; -
CC EMBL; M84991; AAA23738.1; -
CC EMBL; AF000210; AAC74176.1; -
CC EMBL; D90745; BAA35900.1; -
CC PIR; S20443; S20443.
CC PIR; B41856; B41856.
CC PDB; 1MLA; 25-JAN-95.
CC SWISS-2DPAGE; P23715; COLI.
CC EcoGene; EGI1317; fabD.
CC InterPro; IPR001227; AC transferase.
CC InterPro; IPR004410; FabD.
CC Pfam; PF00698; Acyl transf; 1.
CC TIGRFAMs; TIGR00128; fabD; 1.
CC Fatty acid biosynthesis; Transferase; 3D-structure; Complete proteome.
CC INIT MET 0
CC FT ACT SITE 91
CC FT ACT SITE 200
CC FT ACT SITE 200
CC SEQUENCE 308 AA; 32286 MW; F5901043D92PED8E CRC64;
Query Match 42.1%; Score 638.5; DB 1; Length 308;
Best Local Similarity 45.3%; Pred. No. 5e-37;
Matches 140; Conservative 60; Mismatches 102; Indels 7; Gaps 4;
QY 2 TKTALFAGCGAQYILGMGRDFDYQYPIVKEITIDRASQVLYGLDYLLDTDEE--DKLNQT 58
Db 1 TOPAFVFGQSGQVGMGLADMAASPIVEETFAEASALGYDLWAL--TQGGPAAEELNKT 58
QY 59 RYTPQAILATSVAIYRLLOKGYQ--PDMVAGLSLGEYSALVASGALDFEDAVALKRGA 117
Db 59 WOTQPALLTASVALYRVNQVQGGKAPAMWAGHSLGEYSALVCAGVIDFADAVLVNRGK 118
QY 118 YNEEAPADSGKRWAVLNTPTVEIEEACQKASELGVTTPANTYPAQIVIAGEVAVDRA 177
Db 119 FQNEAVPEGTGAAVIAIGLDDASIAKACEEAAEGGVSPVNFNSPGQVVIAGHKEAVERA 178
QY 178 VELLOEAGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVSFSDPTCPVGNTEAAVMQK 237
Db 179 GAACAAAGAKRALPLVSPVSHCALMKPAADKLAELAKITFNAPTVPVNVNVDKCEIN 238
QY 238 ED-IAQLLTROVKEPVRFVSEGVNQEAGISNFIIGPKVLGKVFVKIDTALHAVED 296
Db 239 GDAIRDALVRLQYNPVQWTKSVYMAAQGVHELVYEGPKVLTGLTKRIVDTLTASALINE 299
QY 297 QASLVALLE 305
Db 299 PSAMRAALE 307

Db 181 AALCKEAGAKRALPLAVSVFHCALMKPAAEQALAVLENTQINTPTISVNNVDVKAETE 240
 Qy 237 KEDIAQLLTQVKEPVRFYESIGVMQAGISNFIETGPGKVLGSGFVKKIDQTAHLAHVED 296
 Db 241 GTEIRTAIVQLVSPVETVEKMAQDGVILVLAEVGPGKVLNGLTKRIVGDLQAIISVND 300
 Qy 297 QASLVALLE 305
 Db 301 VASFNAVEE 309

RESULT 4

FABD_SALTY STANDARD; PRT; 308 AA.
 AC 085140;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl-CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
 GN FABD OR STM1194;
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 CX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L72;
 RX MEDLINE=99317265; PubMed=9642179;
 RA Zhang Y., Cronan J.E. Jr.;
 RT "Transcriptional analysis of essential genes of the Escherichia coli
 fatty acid biosynthesis gene cluster by functional replacement with
 the analogous Salmonella typhimurium gene cluster.";
 J. Bacteriol. 180:3295-3303(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 CC -!- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 malonyl-[acyl-carrier protein].
 CC -!- PATHWAY: Fatty acid biosynthesis.
 CC -!- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 TRANSACYLASE ACTIVITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; AF044668; AAC38649.1; -;
 CC EMBL; AE008752; AAL20123.1; -;
 CC HSSP; P25715; 1MLA;
 CC StyGene; SG1637; fabd.
 DR InterPro; IPR001227; AC transferase.
 DR InterPro; IPR004410; FabD.
 DR Pfam; PF00698; Acyl_transf. 1.
 DR TIGRFAMs; TIGR00128; fabd; 1.
 KW Fatty acid biosynthesis; Transferase; Complete proteome.
 FT INIT MET 0
 FT ACT_SITE 91 91 BY SIMILARITY.
 FT ACT_SITE 200 200 BY SIMILARITY.
 SQ SEQUENCE 308 AA; 32274 MW; B3565B545341A611A CRC64;

Query Match 40.3%; Score 611.5; DB 1; Length 308;
 Best Local Similarity 44.1%; Pred. No. 3.5e-35;
 Matches 138; Conservative 62; Mismatches 96; Indels 17; Gaps 5;
 Qy 2 TKTAFLFAGGAQVLGMRDFFDYDPIVKETIDRASQVLGDLRLYLIDTBE---DKLNQT 58
 Db 1 TQPAFVFGGQSGVGLAEMANYPIVETFAESAALGYDLWAL--TOGGPAEELNKT 58
 Qy 59 RYTOPAILATSVAIYRLLOEK-GYQPDWVAGLSLGEYSALVASGALDFEDAVVALVAKGA 117
 Db 59 WQTOPALLTASVALRVWVQGGKVPALMAGHSLSGEYSALVCAGVINFADAVRIVEVRGK 118
 Qy 118 YMEEAAPADSGKVVAVLNTPVEVIEEACOKASELGVVTPANYNTPAQIVTAGEVVAVDRA 177
 Db 119 FMQEAPEGTGMSAIIIGLDDASIAKACBSAEQCVSPVNFNSPGQVVIAGHKEAVERA 178
 Qy 178 VELLQEAQAKELIPLKVGSPHTALLPASOKLAETLAQVSFSDFTCLVGN-----TE 231
 Db 179 GAACKAAGAKRALPLVPVSPHSCALMKFAADKLAVELAKITFSAPTVPVNVNVDVKCETD 238
 Qy 232 AAVQCKEDIAQLLTROVKRPFYFYESIGVMQAGISNFIETGPGKVLGSGFVKKIDQTAHL 291
 Db 239 AAA-----IRDALVRQLYNPQWTKSVEFTAAQGVHELYEVGPGKVLITGLTKRIVDTLTA 293
 Qy 292 AHVEDQASLVALL 304
 Db 294 SALNEPAALSAA 306

RESULT 5

FABD_SINY3 STANDARD; PRT; 293 AA.
 AC P73242;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Malonyl-CoA-acyl carrier protein transacylase (EC 2.3.1.39) (MCT).
 GN FABD OR SLR2023.
 OS Synecocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 CX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsumoto A., Muraki N., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synecocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 DNA Res. 3:109-136(1996).
 CC -!- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 malonyl-[acyl-carrier protein].
 CC -!- PATHWAY: Fatty acid biosynthesis.
 CC -!- SIMILARITY: SIGNIFICANT, TO SEVERAL OTHER PROTEINS WITH
 TRANSACYLASE ACTIVITY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; D90904; BAA17269.1; -;
 CC HSSP; P25715; 1MLA;
 DR InterPro; IPR001227; Ac transferase.
 DR InterPro; IPR004410; FabD.
 DR Pfam; PF00698; Acyl_transf. 1.
 DR TIGRFAMs; TIGR00128; fabd; 1.

KW Paddy acid biosynthesis; Transferase; Complete proteome.
 FT ACT SITE 88 BY SIMILARITY.
 FT ACT SITE 188 BY SIMILARITY.
 SQ SEQUENCE 293 AA; 31491 MW; 31B4ABA59ECDB377 CRC64;

Query Match 30.2%; Score 459; DB 1; Length 293;
 Best Local Similarity 36.8%; Pred. No. 9.8e-25;
 Matches 111; Conservative 59; Mismatches 118; Indels 14; Gaps 6;

QY 3 KTAFLFAGCAQYGLMGDRDFYDQVPIVKETIDRASQVLYGLRYLIDTEEDKLNQRYTQ 62
 Db 2 KTAWFGGCGAAGVGGDILLST-AIAKEYQQAEEILGWSWEKCGQDEASALTQNTQ 60
 QY 63 PAILLATSVAILLQKQYQDMVAGLSGEYSALVAGSALDGEDAVALVAKRGAYVEEA 122
 Db 61 PCLYVIEAILADLLKQGFQDYVAGHSLGEYSALYAGVDFATGQLVKQSEVM--- 117
 QY 123 APADSGKMAVLNTEVEVIEACQKASELGVVVTNPANTPAQIVIAGEVAVDRAVELLQ 182
 Db 118 ASASGMMMAALMKPDQTQLQALTDNTE---VVLANDNSPQVVISGTVAGVE---AILA 171
 QY 183 EGAARLPLKVGSPFHALLLEPASOKLAETLAQVSDFTCPVLGNTEAAVMQKED-IA 241
 Db 172 NVKARVPLKVGSPFHALLLEPASOKLAETLAQVSDFTCPVLGNTEAAVMQKED-IA 241
 QY 242 QLLTQVKEPVRVFSISGMOEAGISNFIETGPGKVLGSGFKVIDQTAHLAHV---EDQA 298
 Db 232 EKLIQMTGVSVRWRTWNLGEIGATDYWEVPGKVLGSGFKVIDQTAHLAHV---EDQA 298
 QY 299 SL 300
 Db 292 SL 293

RESULT 6
 ERYL_SACER STANDARD; PRT; 3491 AA.
 ID ERYL_SACER STANDARD; PRT; 3491 AA.
 AC Q03131;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Erythronolide synthase, modules 1 and 2 (EC 2.3.1.94) (ORF 1) (6-
 DE deoxyerythronolide B synthase I) (DESS 1).
 GN ERYA.
 OS Saccharopolyspora erythraea (Streptomyces erythraeus).
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
 OC Saccharopolyspora.
 OX NCBI_TaxID=1836;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220065; PubMed=2024119;
 RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
 RT "Modular organization of genes required for complex polyketide
 biosynthesis.";
 RL Science 252:675-679 (1991).
 RN [2]
 RP SEQUENCE OF 3474-3491 FROM N.A.
 RX MEDLINE=93231529; PubMed=8386127;
 RA Donadio S., Staver M.J.;
 RT "IS1136, an insertion element in the erythromycin gene cluster of
 Saccharopolyspora erythraea.";
 RL Gene 126:147-151 (1993).
 CC -!- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
 CC deoxyerythronolide B.
 CC -!- COFACTOR: NADP; CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTHETINES.
 CC -!- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
 CC BIOSYNTHESIS.
 CC -!- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
 CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
 CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
 CC SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS
 CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
 AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
 RESPECTIVELY.
 CC -!- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
 CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
 CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
 CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
 CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
 CC OF THE FULL-LENGTH CHAIN.
 CC -!- SIMILARITY: TO PATTY ACID SYNTHASE (FAS).
 CC -!- SIMILARITY: CONTAINS 3 ACYL CARRIER DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; MG3676; AAA26493.2; -;
 DR EMBL; L07626; AAA26504.1; -;
 DR HSSP; P25715; 1MLA.
 DR InterPro; IPR002128; ADH short.
 DR InterPro; IPR001227; AC_Transferase.
 DR InterPro; IPR000794; Ketoacyl-synt.
 DR InterPro; IPR003880; Ppantne attach.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00550; pp-binding; 3.
 DR Pfam; PF00898; Acyl_transf; 3.
 DR Pfam; PF02801; ketoacyl-synt_C; 2.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 3.
 DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
 DR PROSITE; PS00075; ACP DOMAIN; 3.
 KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
 KW Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1972 MODULE 1.
 FT DOMAIN 1979 3491 MODULE 2.
 FT DOMAIN 1 375 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 414 484 ACYL CARRIER (ACP) 1.
 FT DOMAIN 503 961 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 1030 1356 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 1611 1794 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1888 1958 ACYL CARRIER (ACP) 2.
 FT DOMAIN 1979 2441 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2507 2854 ACYLTRANSFERASE (AT) 3.
 FT DOMAIN 3055 3237 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 3334 3404 ACYL CARRIER (ACP) 3.
 FT ACT SITE 145 145 ACYL-ENZYME INTERMEDIATE.
 FT BINDING 447 447 PHOSPHOPANTHETINE (BY SIMILARITY).
 FT ACT_SITE 677 677 BETA-KETOACYL SYNTHASE.
 FT ACT_SITE 1128 1128 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 1614 1660 NADP.
 FT BINDING 1921 1921 PHOSPHOPANTHETINE (BY SIMILARITY).
 FT ACT_SITE 2148 2148 BETA-KETOACYL SYNTHASE.
 FT ACT_SITE 2598 2598 ACYL-ENZYME INTERMEDIATE.
 FT NP_BIND 3058 3104 NADP.
 FT BINDING 3367 3367 PHOSPHOPANTHETINE (BY SIMILARITY).
 SQ SEQUENCE 3491 AA; 365022 MW; 652BFC32C90FA8C4 CRC64;

Query Match 20.1%; Score 305.5; DB 1; Length 3491;
 Best Local Similarity 29.2%; Pred. No. 7.7e-13;
 Matches 93; Conservative 58; Mismatches 134; Indels 33; Gaps 10;

QY 4 TAFLLAGCAQYGLMGDRDFYDQVPIVKETIDRA-----SOVLGYDLRYLIDTEED--KLNQ 57
 Db 2507 SVVFFGQAGQAGMGRDFYDQVPIVKETIDRA-----SOVLGYDLRYLIDTEED--KLNQ 57
 QY 58 TRYQPAILLATSVAILRLLQKQYQDMVAGLSGEYSALVAGSALDGEDAVALVAKRGA 117
 Db 2566 VDVVQVFLVAVVYSLARLWACGAVPSAVIGHSGQELAAVAVAGALSLEDGMEVVARSR 2625
 QY 118 YMEERAPADSGKMAVLNTEVEVIEACQKASELGVVVTNPANTPAQIVIAGEVAVDRA 177

```

DR Tuberculin; RV2243; -.
DR InterPro: IPR001227; Acyl transferase.
DR Pfam: PF00698; Acyl transf, 1.
KW Fatty acid biosynthesis; Transferase; Complete proteome.
FT ACT_SITE 91 91 BY SIMILARITY.
FT ACT_SITE 194 194 BY SIMILARITY.
FT DOMAIN 206 210 POLY-ALA.
FT DOMAIN 233 236 POLY-ALA.
SQ SEQUENCE 302 AA; 30788 MW; BB7BCD8217FC66C8 CRC64;

Query Match 18.1%; Score 274.5; DB 1; Length 302;
Best Local Similarity 29.9%; Pred. No. 4.8e-12;
Matches 92; Conservative 52; Mismatches 135; Indels 29; Gaps 9;

QY 5 AFLFAGGAGYLGMRDFYDQYFIVKETTDRASQVGLYDLRYLIDT-EEDKLNQTRYTP 63
DB 3 ALLAPGQSGTQTEGMLSPWL-QLFCAADQIAAWKAADLDLARLGTASTEEITDTVAQP 61

QY 64 AILATSVAIYRLLEK- --GYQPMVAGLSLGEYSALVASGALDFEDAVALKRGAYME 120
DB 62 LIVATILLAHQELARRCVLACKDVIIVAGSHVGIAYAYAGVIAADDAVALAATRGEMA 121

QY 121 EAAPADSGKVMVINTVEVIEBACQKASLGVTVPANNYTPAQIVIAGEVAVDRAVEL 180
DB 122 KACATEPTGMSAVLGDE- --TEVLSRLQLDLV-PANRNAGQIVAAAGRLTALEK- --L 174

QY 181 LOEAGAK-RLIPLVKVGPPHTALLEPASOKLAETLAQVSFSDFTCLVGNTEAAVMQKED 239
DB 175 AEDPAPKARVALGVAGAFHTEFMAPALDGAFAAANIANIATADPTATLLSNRDKPVTSA 234

QY 240 IA-QLLTRQVKEPVRFYESIGVMQEAGISNFIIGPKVLGPFVK- ----- 284
DB 235 AAMDVLVSQLTQPVWMDLTATLREHTVTAIVEFPFAGTSLGIAKRELGVPARAVKSPA 294

QY 285 -IDQTAHL 291
DB 295 DLDELALN 302

RESULT 8
STCA_EMENI STANDARD; PRT; 2181 AA.
ID STCA_EMENI
AC Q12397;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative sterigmatocystin biosynthesis polyketide synthase (PKS).
GN STCA OR PKSST.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 26;
RX MEDLINE=98202293; PubMed=8643646;
RA Brown D.W., Yu J.-H., Kellar H.S., Fernandes M., Nesbitt T.C.,
RA Kellar N.P., Adams T.H., Leonard T.J.;
RT "Twenty-five coregulated transcripts define a sterigmatocystin gene
RT cluster in Aspergillus nidulans."
RL Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4;
RX MEDLINE=95370159; PubMed=7642507;
RA Yu J.-H., Leonard T.J.;
RT "Sterigmatocystin biosynthesis in Aspergillus nidulans requires a
RT novel type I polyketide synthase."
RL J. Bacteriol. 177:4792-4800 (1995).
CC -!- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
CC STERIGMATOCYSTIN FROM HEXANOYL-COA AND SEVEN MALONATES.
CC -!- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETINES
CC (POTENTIAL).

```

```

DR Tuberculin; RV2243; -.
DR InterPro: IPR001227; Acyl transferase.
DR Pfam: PF00698; Acyl transf, 1.
KW Fatty acid biosynthesis; Transferase; Complete proteome.
FT ACT_SITE 91 91 BY SIMILARITY.
FT ACT_SITE 194 194 BY SIMILARITY.
FT DOMAIN 206 210 POLY-ALA.
FT DOMAIN 233 236 POLY-ALA.
SQ SEQUENCE 302 AA; 30788 MW; BB7BCD8217FC66C8 CRC64;

Query Match 18.1%; Score 274.5; DB 1; Length 302;
Best Local Similarity 29.9%; Pred. No. 4.8e-12;
Matches 92; Conservative 52; Mismatches 135; Indels 29; Gaps 9;

QY 5 AFLFAGGAGYLGMRDFYDQYFIVKETTDRASQVGLYDLRYLIDT-EEDKLNQTRYTP 63
DB 3 ALLAPGQSGTQTEGMLSPWL-QLFCAADQIAAWKAADLDLARLGTASTEEITDTVAQP 61

QY 64 AILATSVAIYRLLEK- --GYQPMVAGLSLGEYSALVASGALDFEDAVALKRGAYME 120
DB 62 LIVATILLAHQELARRCVLACKDVIIVAGSHVGIAYAYAGVIAADDAVALAATRGEMA 121

QY 121 EAAPADSGKVMVINTVEVIEBACQKASLGVTVPANNYTPAQIVIAGEVAVDRAVEL 180
DB 122 KACATEPTGMSAVLGDE- --TEVLSRLQLDLV-PANRNAGQIVAAAGRLTALEK- --L 174

QY 181 LOEAGAK-RLIPLVKVGPPHTALLEPASOKLAETLAQVSFSDFTCLVGNTEAAVMQKED 239
DB 175 AEDPAPKARVALGVAGAFHTEFMAPALDGAFAAANIANIATADPTATLLSNRDKPVTSA 234

QY 240 IA-QLLTRQVKEPVRFYESIGVMQEAGISNFIIGPKVLGPFVK- ----- 284
DB 235 AAMDVLVSQLTQPVWMDLTATLREHTVTAIVEFPFAGTSLGIAKRELGVPARAVKSPA 294

QY 285 -IDQTAHL 291
DB 295 DLDELALN 302

RESULT 8
STCA_EMENI STANDARD; PRT; 2181 AA.
ID STCA_EMENI
AC Q12397;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative sterigmatocystin biosynthesis polyketide synthase (PKS).
GN STCA OR PKSST.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 26;
RX MEDLINE=98202293; PubMed=8643646;
RA Brown D.W., Yu J.-H., Kellar H.S., Fernandes M., Nesbitt T.C.,
RA Kellar N.P., Adams T.H., Leonard T.J.;
RT "Twenty-five coregulated transcripts define a sterigmatocystin gene
RT cluster in Aspergillus nidulans."
RL Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FGSC 4;
RX MEDLINE=95370159; PubMed=7642507;
RA Yu J.-H., Leonard T.J.;
RT "Sterigmatocystin biosynthesis in Aspergillus nidulans requires a
RT novel type I polyketide synthase."
RL J. Bacteriol. 177:4792-4800 (1995).
CC -!- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF
CC STERIGMATOCYSTIN FROM HEXANOYL-COA AND SEVEN MALONATES.
CC -!- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETINES
CC (POTENTIAL).

```



```
Db 1085 RACSEKGLKSTI-LTVYAFHSAQVBEILEDLQALQGITFNKPSVPFVSALLGEVITEA 1143
QY 239 DI-----AQLLTROVKPEVRVESIGVMQEAAGISN-----FIEIGPGKVLGSGFVK 283
Db 1144 GSNILNAEYLVHRCRETNVNLSFAFEAVRNAKLGGDQTLMLVGPHTVCSGMVK 1196

RESULT 10
PKSI ASPPA
ID PKSI ASPPA STANDARD; PRT; 2109 AA.
AC Q12053;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aflatoxin biosynthesis polyketide synthase (PKS).
GN PKSL1.
OS Aspergillus parasiticus.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2999;
RX MEDLINE=96042102; PubMed=7592391;
RA Feng G.H., Leonard T.J.;
RT "Characterization of the polyketide synthase gene (pkSL1) required for aflatoxin biosynthesis in Aspergillus parasiticus."
RL J. Bacteriol. 177:6246-6254(1995).
CC -!- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE POLYKETIDE NUCLEUS OF AFLATOXIN FROM HEXANOYL COA AND SEVEN MALONATES.
CC -!- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTETHEINE (BY SIMILARITY).
CC -!- PATHWAY: Aflatoxin biosynthesis; first step.
CC -!- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L42766; AAC41675.1; -.
DR EMBL; L42765; AAC41674.1; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Peptide attach.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF00550; pp-binding; 1.
DR Pfam; PF00698; Acyl trans; 1.
DR Pfam; PF00975; Thioesterase; 1.
DR Pfam; PF02801; ketoacyl-synt C; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
DR TRANSFERASE; Acyltransferase; Phosphopantetheine;
KW Multifunctional enzyme.
FT DOMAIN 374 805 BETA-KETOACYL SYNTHASE.
FT DOMAIN ? ? ACYL/MALONYL TRANSFERASES.
FT DOMAIN 1714 1785 ACYL CARRIER (ACP).
FT DOMAIN ? 2109 THIOESTERASE.
FT ACT_SITE 543 543 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 993 993 ACYL/MALONYL TRANSFERASES (BY SIMILARITY).
FT BINDING 1746 1746 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT_SITE 1937 1937 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 2109 AA; 230715 MW; CB701372A16D8551 CRC64;

Query Match 16.7%; Score 254; DB 1; Length 2109;
Best Local Similarity 28.9%; Pred. No. 1.4e-09;
Matches 87; Conservative 48; Mismatches 132; Indels 34; Gaps 11;
-QY 3 KTAFLAQGQAQYLGMGRDFYDQYPIVKETI---DRASQVLGYDLRYLI-----DTEED 53
```

```
Db 900 KILFAFTGGQSYATMGQKVYDAYSPFREDLEKFDRLAQSHPFSLHVCTSPKGDVEE- 958
QY 54 KLNQTRVYTPAILATSVAIVRLLOEKGYQYQDMVAGLSLGRYSALVAGSALDFFEDAVLVA 113
Db 959 --MAPVVVQLAITCLQMALTNLMTSFGIRPDVTVGHSLGFEAALYAAGVLSASDVVYLVG 1016
QY 114 KRGAYMEEAPADSGKMVAUNLTPVEVIEACQASLGVVTPANYTPAQIVTAGEVVA 173
Db 1017 QRABELQERCQRGTHAMLVAKATPEALSOWIQDHCDEVACI-----NGPEDTVLSGTTKN 1071
QY 174 VDRAVELLQEAQAKRLIPLKVGSPFHTALLEPASQKLAETLAQ-VSFSD-----FTCPVLVG 228
Db 1072 VAEVQRAMTDNGIKCTL-LKLPFAFSAQVQPIILDDF-EALAQAQATFAKQLLILSPLLR 1129
QY 229 NT--EAAVMQKEDIAQLLTROVKPEVRVESIGVMQEAAGISN-----FIEIGPGKVLGSGFV 282
Db 1130 TEIHQGVVTPSYVAQ-----HCRHTVDMQAALRSAREKGLIDDKTLVIELGPKPLISGMV 1185
QY 283 K 283
Db 1186 K 1186

RESULT 11
OL56 STRAT
ID OL56 STRAT STANDARD; PRT; 3519 AA.
AC Q07017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oleandomycin polyketide synthase, modules 5 and 6.
GN ORPB.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150470; PubMed=8107683;
RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
RT "Characterisation of a Streptomyces antibioticus gene encoding a type I polyketide synthase which has an unusual coding sequence."
RL Mol. Gen. Genet. 242:358-362(1994).
CC -!- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN LACTONE RING.
CC -!- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
CC -!- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L09654; AAA19695.1; -.
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Peptide attach.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00109; ketoacyl-synt; 2.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00698; Acyl trans; 2.
DR Pfam; PF00975; Thioesterase; 1.
DR Pfam; PF02801; ketoacyl-synt C; 2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
DR PROSITE; PS50075; ACP_DOMAIN; 2.
KW Phosphopantetheine; Multifunctional enzyme; Repeat.
KW
```



```

FT DOMAIN 1 2
FT DOMAIN 32 3519
FT DOMAIN 569 890
FT DOMAIN 1200 1382
FT DOMAIN 1487 1561
FT DOMAIN 1686 2156
FT DOMAIN 2230 2541
FT DOMAIN 2856 3038
FT DOMAIN 3141 3215
FT ACT SITE 1 2
FT BINDING 3219 3519
FT DOMAIN 210 210
FT ACT SITE 660 660
FT NP_BIND 1203 1249
FT BINDING 1524 1594
FT ACT SITE 1859 1859
FT ACT SITE 2311 2311
FT NP_BIND 2859 2905
FT BINDING 3178 3178
SQ SEQUENCE 3519 AA; 368561 MW; 41AE78AAAE61F86 CRC64;

Query Match 16.7%; Score 253.5; DB 1; Length 3519;
Best Local Similarity 26.2%; Pred. No. 2.9e-09;
Matches 84; Conservative 61; Mismatches 137; Indels 39; Gaps 10;

QY 3 KTAFLFAGGAGYLGMRDFFDQYPIVKTIDRASQVL---GYDL-RYLIDTDEEDKLNQ 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2219 RVVFPGGSGWVGAGLLDACVFPAEVAECAAVLDPVTGWSLVEVLQGRDATVLGR 2278

QY 58 TRYTPAILATSVAIRYLQEKGYQDMVAGLSLGEYSALVAGSALDPED----- 107
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2279 VDWQPALVAMVSLARTWRYGVFEPAVAVGHSQGEIAAACVAGSLADGARVWVLRSR 2338

QY 108 AVALAKGAYNEEAAPADSGWAVLNPTEVIEEACQKASELGVTPANNYTPAQIVI 167
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2339 AIARAGGGWVSVSLPA--GVRVRLDTP-----YGRVSVAAVNGPSSFTV 2383

QY 168 AGEVAVDRAVELLQEGAGAK-RLIPLKVGSPFHTALLBPASQKLAETLAQVSFSDTCPL 226
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2384 SGDVALDELLAGCEREGVRARRVPVDYAS--HSAQMDQLRDELLEALDITPDQSSVFF 2441

QY 227 VQNTAAVWQKEDI-AQLLTROKPEVPFYSIGVWQAGISNTEIGQKVLGSGFYKKI 285
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2442 FSTVTADMLDTTALDAGVWFTNLRETVRFQAEVGLVAGQMGAFVCECSFHPVL---VPGI 2498

QY 286 DQTAHLAHVEDQASVVALLEK 306
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2499 EGT--LDALDQNAVLGSLRR 2517

RESULT 12
ERY2 SACER STANDARD; PRT; 3567 AA.
AC Q03132; Q54096;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 3 and 4 (EC 2.3.1.94) (ORF 2) (6-
DE deoxyerythronolide B synthase II) (DEBS 2).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteria (Class); Actinobacteridae;
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Scaver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide
biosynthesis."

```

```

RL Science 252:675-679 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92155230; PubMed=1740151;
RA Revitt D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.
RT Cloning of the structural gene, sequence analysis and inferred domain
structure of the multifunctional enzyme."
RL Eur. J. Biochem. 204:39-49 (1992).
CC -1- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -1- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHETHEINES.
CC -1- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -1- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -1- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -1- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M63677; AAA36494.1; -
CC EMBL; X62569; CAA44448.1; -
CC InterPro; IPR001227; Ac transferase.
CC InterPro; IPR002085; Adh_zn_family.
CC InterPro; IPR000794; Ketoacyl-synt.
CC InterPro; IPR003880; Ppantne_attach.
CC Pfam; PF00109; adh_zinc; 1.
CC Pfam; PF00109; ketoacyl-synt; 2.
CC Pfam; PF00550; pp-binding; 2.
CC Pfam; PF00698; Acyl_transf; 2.
CC Pfam; PF02801; ketoacyl-synt_C; 2.
CC PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
CC PROSITE; PS00606; B_KETOACYL SYNTHASE; 2.
CC PROSITE; PS50075; ACP DOMAIN; 2.
CC Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
CC Phosphopantetheine; Multifunctional enzyme.
CC FT DOMAIN 1 1484 MODULE 3.
CC FT DOMAIN 1485 3567 MODULE 4.
CC FT DOMAIN 27 489 BETA-KETOACYL SYNTHASE 1.
CC FT DOMAIN 559 884 ACYLTRANSFERASE (AT) 1.
CC FT DOMAIN 1130 1301 BETA-KETOACYL REDUCTASE 1 (POSSIBLY
CC NON-FUNCTIONAL).
CC FT DOMAIN 1397 1467 ACYL CARRIER (ACP) 1.
CC FT DOMAIN 1485 1943 BETA-KETOACYL SYNTHASE 2.
CC FT DOMAIN 2013 2336 ACYLTRANSFERASE (AT) 2.
CC FT DOMAIN 2383 3066 DEHYDRATASE/ENOYLREDUCTASE (DH/ER).
CC FT DOMAIN 3139 3322 BETA-KETOACYL REDUCTASE 2.
CC FT DOMAIN 3415 3485 ACYL CARRIER (ACP) 2.
CC FT ACT SITE 202 202 THIOESTER BOND.
CC FT ACT SITE 651 651 ACYL-ENZYME INTERMEDIATE.
CC FT BINDING 1430 1430 PHOSPHOPANTHETHEINE (BY SIMILARITY).
CC FT ACT_SITE 1661 1661 THIOESTER BOND.
CC FT ACT_SITE 2115 2115 ACYL-ENZYME INTERMEDIATE.
CC FT NP_BIND 2961 2961 NADP (ER).

```

BEST

```

FT NP_BIND 3142 3157 NADP (KR).
FT BINDING 3448 3448 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT CONFLICT 438 438 R -> A (IN REF. 2).
FT CONFLICT 480 480 T -> S (IN REF. 2).
FT CONFLICT 1241 1241 L -> F (IN REF. 2).
FT CONFLICT 2664 2664 G -> V (IN REF. 2).
SQ SEQUENCE 3567 AA; 374413 MW; E86284F4738AA0C0 CRC64;

Query Match 16.24; Score 246; DB 1; Length 3567;
Best Local Similarity 26.5%; Pred. No. 9.6e-09;
Matches 80; Conservative 53; Mismatches 149; Indels 20; Gaps 8;

QY 3 KTAFLPAGQAOYLGMGRDQYQYPIVKETIDRASQVLGVOLRYLI-----DTEEDKLN 56
DB 2012 KPVLPFGQAGWGMARDLLESEVPASMSCAELSPHDTDKLVDVRGGDPDPE 2071

QY 57 QTRYTPAILATSAIVRLIQXYQYQDMVAGLSLGEYSALVASGALDFEDAVLVAKRG 116
DB 2072 RVDVLQVLFPSIMVSLAELWRAHGVTTPAAVVGHSQGEIAAAHVAGALSLEAAKVVALS 2131

QY 117 AYMEERAPADSGKMAVINTPVEIIEACOKASELGVTPANNVTQAQIVIAGEVVADR 176
DB 2132 QVREL--DQGGWVSV-GASRDELETVLARWD--GRVAVAVNGPSTSVAGPTAELD- 2185

QY 177 AVELLOPAGAKRLIPKVSQGF--HTALLEPASQKLAETIAQVSFSDFTCLPLVGNTEAAV 234
DB 2186 --EPPAEAREMKPRIAVRYASHSPVARIEDRLAELGTITAVRGSVPLHSTVTGGEV 2243

QY 235 MQKEDI-AQLLTQVKEPVRFYSIGVMQFAGISNFIETGPGKVLGFKKIDQTAHLAH 293
DB 2244 IDTSAMDASWYWRNLRRPVLFEQAVRGLVEQGFDTFVEVSPHPVL---LMAVEETAHAG 2300

QY 294 VE 295
DB 2301 AE 2302

RESULT 13
ERY3 SACER STANDARD; PRT; 3172 AA.
AC Q0133; O54097; Q99270;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Erythronolide synthase, modules 5 and 6 (EC 2.3.1.94) (ORF 3) (6-
DE deoxyerythronolide B synthase III) (DEBS 3).
GN ERYA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Pseudonocardiaceae; Pseudonocardiaceae;
OC Saccharopolyspora.
OX NCBI_TaxID=1836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=91043075; PubMed=2234082;
RA Cortes J., Haydock S.F., Roberts G.A., Bevitt D.J., Leadlay P.F.;
RT "An unusually large multifunctional polypeptide in the erythromycin-
RL producing polyketide synthase of Saccharopolyspora erythraea.";
RN Nature 348:176-178(1990).
RP SEQUENCE FROM N.A.
RX MEDLINE=91220065; PubMed=2024119;
RA Donadio S., Staver M.J., McAlpine J.B., Swanson S.J., Katz L.;
RT "Modular organization of genes required for complex polyketide
RL biosynthesis.";
RN Science 252:675-679(1991).
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL 2338;
RX MEDLINE=92145230; PubMed=1740151;
RA Bevitt D.J., Cortes J., Haydock S.F., Leadlay P.F.;
RT "6-Deoxyerythronolide-B synthase 2 from Saccharopolyspora erythraea.

```

```

RT Cloning of the structural gene, sequence analysis and inferred domain
RT structure of the multifunctional enzyme.";
RL Eur. J. Biochem. 204:39-49(1992).
CC -|- CATALYTIC ACTIVITY: 6 malonyl-CoA + propionyl-CoA = 7 CoA + 6-
CC deoxyerythronolide B.
CC -|- COFACTOR: NADP; CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES.
CC -|- PATHWAY: COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN
CC BIOSYNTHESIS.
CC -|- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH
CC ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3
CC ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH
CC SYNTHASE PARTICIPATES IN ONE OF THE SIX PAS-LIKE ELONGATION STEPS
CC REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5,
CC AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6,
CC RESPECTIVELY.
CC -|- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT),
CC BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER
CC PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETO REDUCTASE (KR),
CC DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE
CC BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION
CC OF THE FULL-LENGTH CHAIN.
CC -|- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).
CC -|- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X56107; CAA39583.1; -.
CC EMBL; M63677; AAA26495.1; -.
CC EMBL; X62569; CAA44449.1; -.
CC HSPF; P00101; ICH.
CC InterPro; IPR002198; ADH short.
CC InterPro; IPR001227; Ac transferase.
CC InterPro; IPR000794; Ketoacyl-synt.
CC InterPro; IPR003880; Pantine attach.
CC InterPro; IPR001031; Thioesterase.
CC Pfam; PF00105; adh short; 1.
CC Pfam; PF00109; ketoacyl-synt; 2.
CC Pfam; PF00550; pp-binding; 2.
CC Pfam; PF00698; Acyl trans; 2.
CC Pfam; PF00975; Thioesterase; 1.
CC Pfam; PF02801; ketoacyl-synt C; 2.
CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
CC PROSITE; PS00606; B KETOACYL SYNTHASE; 2.
CC PROSITE; PS50075; ACP_DOMAIN; 2.
KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat;
KW Phosphopantetheine; Multifunctional enzyme.
FT DOMAIN 1 1484 MODULE 5.
FT DOMAIN 1485 3172 MODULE 6.
FT DOMAIN 37 484 BETA-KETOACYL SYNTHASE 1.
FT DOMAIN 554 878 ACYLTRANSFERASE (AT) 1.
FT DOMAIN 1116 1298 BETA-KETOACYL REDUCTASE 1.
FT DOMAIN 1394 1464 ACYL CARRIER (ACP) 1.
FT DOMAIN 1488 1954 BETA-KETOACYL SYNTHASE 2.
FT DOMAIN 2021 2335 ACYLTRANSFERASE (AT) 2.
FT DOMAIN 2555 2735 BETA-KETOACYL REDUCTASE 2.
FT DOMAIN 2821 2891 ACYL CARRIER (ACP) 2.
FT DOMAIN 2926 3172 THIOESTERASE.
FT ACT_SITE 199 199 THIOESTER BOND.
FT ACT_SITE 643 643 ACYL-ENZYME INTERMEDIATE.
FT NP_BIND 1118 1164 NADP.
FT BINDING 1427 1427 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT_SITE 1661 1661 THIOESTER BOND.
FT ACT_SITE 2112 2112 ACYL-ENZYME INTERMEDIATE.
FT NP_BIND 2557 2605 NADP.
FT BINDING 2854 2854 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT CONFLICT 231 231 S -> SA (IN REF. 2).
FT CONFLICT 240 240 MISSING (IN REF. 2).

```

```
FT CONFLICT 289 A -> R (IN REF. 2).
FT CONFLICT 289 P -> R (IN REF. 2).
FT CONFLICT 493 PEPRNSLRDTGTLATRASAMEHRA -> ASRGTRCATPVS
FT CONFLICT 517 RWPAAADWEO (IN REF. 1).
FT CONFLICT 510 A -> R (IN REF. 2).
FT CONFLICT 513 M -> W (IN REF. 2).
FT CONFLICT 525 E -> D (IN REF. 2).
FT CONFLICT 536 R -> G (IN REF. 2).
FT CONFLICT 547 GNPS -> ATR (IN REF. 2).
FT CONFLICT 553 R -> A (IN REF. 2).
FT CONFLICT 553 R -> A (IN REF. 2).
FT CONFLICT 673 MISSING (IN REF. 2).
FT CONFLICT 716 AHK -> GIT (IN REF. 2).
FT CONFLICT 734 R -> RQR (IN REF. 2).
FT CONFLICT 896 R -> RELVPFORQR (IN REF. 1).
FT CONFLICT 896 R -> RELVPFORQR (IN REF. 1).
FT CONFLICT 994 GVAAYPH -> VLSLRD (IN REF. 2).
FT CONFLICT 1108 RTHPLEA -> ATRWSR (IN REF. 2).
FT CONFLICT 1124 MISSING (IN REF. 1).
FT CONFLICT 1132 L -> V (IN REF. 2).
FT CONFLICT 1132 A -> R (IN REF. 2).
FT CONFLICT 1134 MISSING (IN REF. 2).
FT CONFLICT 1277 AA -> RR (IN REF. 2).
FT CONFLICT 1278 LQDRE -> STAER (IN REF. 2).
FT CONFLICT 1385 MISSING (IN REF. 2).
FT CONFLICT 1485 G -> R (IN REF. 2).
FT CONFLICT 1518 V -> L (IN REF. 2).
FT CONFLICT 1601 LP -> FA (IN REF. 2).
FT CONFLICT 1724 Q -> L (IN REF. 2).
FT CONFLICT 1732 GAEG -> ARRA (IN REF. 2).
FT CONFLICT 1739 T -> S (IN REF. 2).
FT CONFLICT 1762 D -> DGAD (IN REF. 2).
FT CONFLICT 2252 OSP -> AVA (IN REF. 2).
FT CONFLICT 2275 G -> GR (IN REF. 2).
FT CONFLICT 2408 LA -> S (IN REF. 2).
FT CONFLICT 2421 NA -> TH (IN REF. 2).
FT CONFLICT 2444 A -> G (IN REF. 2).
FT CONFLICT 2596 P -> A (IN REF. 2).
FT CONFLICT 2609 RRAEGRA -> AVKAVRR (IN REF. 1).
FT CONFLICT 2715 D -> E (IN REF. 2).
FT CONFLICT 2754 DBBD5094E77DD5F CRC64;
SQ SEQUENCE 3172 AA; 331474 MW; DBBD5094E77DD5F CRC64;

Query Match 16.0%; Score 242.5; DB 1; Length 3172;
Best Local Similarity 29.2%; Pred. No. 1.4e-08;
Matches 84; Conservative 47; Mismatches 134; Indels 23; Gaps 10;

QY 3 KTAFLPAGGAGYLGMRDFDYQYPIVKETIDRASQVLG---YDLRYLIDTEEDKLNQ 58
DB 553 RVAMVFPGGAGQMGQWARDLLRESQVFADSDTCERALAPHVDSLTLLSGAR-PLDRV 611

QY 59 RYTQPAILATSVAILYLLQEKYQPDPMVAGSLGEYSALVASGALDFEDAVLAKRGAY 118
DB 612 DVVQPALFAMVMSLAALWRSHGVEPAAVVGHSGEIAAAHVAGALTLEDAAKLVAVRSR 671

QY 119 MEEAPADSGKAVLNTPEVIEERACOKASEL-GVVTPTANVTP-AQIVIAGEVAVDR 176
DB 672 LARLG-GQGMAFGLGT-----EQAEIRGRFAGALSIVNGRPSVVVAGSGPLDE 725

QY 177 AV-ELLQEAGAKRLIPLKVPSPFHTALLEPASQKLAETLAQVSFDFCTPLVGNTEAAVM 235
DB 726 LIAECEAEAKHARRIPVDVAS--HSPQVESLEELLTELAGISPVSDVALVSTTG---780

QY 236 QKEDTAQLLT---RQVKEPVFYSIGVMQEAAGISNIEIGPGKVL 279
DB 781 QPIDATMDTAYWYANLREQVRFPQATROLAEAGDAFVEVSPHPVLT 828

RESULT 14
PPSB MYCTU
ID_PPSB_MYCTU STANDARD; PRT; 1538 AA.
AC Q10978; OS3234;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
```

```
DE Phenolphthiocerol synthesis polyketide synthase ppsB.
GN PPSB OR RV2932 OR MT3002 OR MTCY338.21 OR MTV011.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jorgels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Rogers R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
CC SYNTHESIS.
CC -1- COFACTOR: CONTAINS 1 COVALENTLY BOUND PHOSPHOPANTHETHEINE (BY
CC similarity).
CC -1- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AL021070; CAA15929.1; --
CC EMBL; AE007122; AAK47329.1; --
CC TIGR; MT3002; --
CC TubercuList; RV2932; --
CC InterPro; IPR001227; Ac transferase.
CC InterPro; IPR00794; Ketoacyl-synt.
CC InterPro; IPR003880; Ppantne attach.
CC Pfam; PF00109; ketoacyl-synt; 1.
CC Pfam; PF00550; PP-binding; 1.
CC Pfam; PF00698; Acyl-transf; 1.
CC Pfam; PF02801; ketoacyl-synt_C; 1.
CC PROSITE; PS00606; B KETOACYL SYNTHASE; 1.
CC PROSITE; PS00012; PHOSPHOPANTHETHEINE; FALSE_NEG.
CC PROSITE; PS00075; ACP DOMAIN; 1.
CC Hypothetical protein; Multifunctional enzyme; Oxidoreductase;
KW Transaminase; NADP; Phosphopantetheine; Complete proteome.
FT DOMAIN 1425 1495 ACYL CARRIER (ACP)
FT BINDING 1458 1458 PHOSPHOPANTHETHEINE (POTENTIAL).
SQ SEQUENCE 1538 AA; 162527 MW; B55E2A2042AD00CC CRC64;

Query Match 14.8%; Score 224; DB 1; Length 1538;
Best Local Similarity 27.0%; Pred. No. 1.1e-07;
Matches 83; Conservative 55; Mismatches 129; Indels 40; Gaps 13;

QY 4 TAFLPAGGAGYLGMRDFDYQYPIVKETIDRASQVL----GYDRLYLIDTEEDKLNQTR 59
DB 560 TVFYVYSGRGSQWAGMGRQLLADEFAAAVALEPVEVQAGSLRDIATGKE-LVGIE 618
```

QY 60 YTOPAILATSVAIYRLLEQKGYQDMVAGLSGLGYSALVAGSALDGFEDAVLVAKRGAYM 119
 Db 619 QIQLGLIGMQLTUTELWRSYGVQPDVLVGHSMGEVAAVAVAGALTPEAGLRVTATRARLM 678
 QY 120 EEARPADSGWAVLNTFVEVIEACOKA--SELGVVTPANTYNTPAQIVIAGEVAVDRA 177
 Db 679 ---APLSGGQGMALLG-----LDAATEALIAIDYPOVTGVIYNSPQTVIAGTQID-- 728
 QY 178 VELLOBAGAKRLIPLKVS---GPFHTAL--LEPASOKLAETLAQVSFSDFT--CPLVG-- 228
 Db 729 -ELIARVRAQNRFPASRVNIEVAPHNPAMDALQAPMES-----ELADLTPTPTIGII 779
 QY 229 NTEAANVQKEDI--AQLLTQVKEPVRFYESI---GVMOEAGISNFIETIGPKVLSGFVK 283
 Db 780 STTYADLHTQIPDAEHWATNMNRPVRFQCAIASAGSGADGAYHTFIEISAPHLTQAIA 839
 QY 284 KIDQTAH 290
 Db 840 DTJEDAH 846

RESULT 15
 PPSS MYCTU STANDARD; PRT; 1876 AA.
 ID PPSS MYCTU
 AC Q10977;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenolphthiocerol synthesis polyketide synthase ppsa.
 GN PPSS OR RV2931 OR MT3000 OR MTCY338.20.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POLYKETIDE SYNTHASE INVOLVED IN PHENOLPTHIOCEROL
 CC SYNTHESIS.
 CC -1- COPACITOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTETHEINES (By
 CC similarity).
 CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

EMBL; Z74697; CAA98988.1; -
 EMBL; AE007122; AAK47328.1; -
 TIGR; MT3000; -
 DR Tuberculist; RV2931; -
 DR Interpro; IPR001227; AC_transferase.
 DR Interpro; IPR00794; Ketoacyl-synt.
 DR Interpro; IPR003880; Pantne_attach.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00698; Acyl trans; 1.
 DR Pfam; PF02801; ketoacyl-synt C; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 DR PROSITE; PS00608; KETOACYL SYNTHASE; 1.
 DR PROSITE; PS00075; ACP DOMAIN; 2.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Repeat;
 KW Phosphopantetheine; Complete proteome.
 FT DOMAIN 7 80
 FT ACYL CARRIER (ACP) 1.
 FT NP_BIND 1764 1833
 FT NADP (POTENTIAL).
 FT BINDING 62 62
 FT PHOSPHOPANTETHEINE (POTENTIAL).
 FT ACT_SITE 273 273
 FT BETA-KETOACYL SYNTHASE (BY SIMILARITY).
 FT ACT_SITE 720 720
 FT MALONYLTRANSFERASE (BY SIMILARITY).
 FT BINDING 1796 1796
 FT PHOSPHOPANTETHEINE (POTENTIAL).
 FT CONFLICT 624 624
 FT D -> E (IN REF. 2).
 FT CONFLICT 877 877
 FT R -> H (IN REF. 2).
 FT CONFLICT 1323 1323
 FT G -> S (IN REF. 2).
 SQ SEQUENCE 1876 AA; 198834 MW; D9783BDB48792110 CRC64;

Query Match 13.7%; Score 208.5; DB 1; Length 1876;
 Best Local Similarity 26.4%; Pred. No. 1.6e-06;
 Matches 81; Conservative 52; Mismatches 135; Indels 39; Gaps 13;

QY 4 TATLFAQGAQYLCMGGRDFYDQYPIVKTIDRASQVL---GYDLRYLIDTEEDKLNQTR 59
 Db 631 TVFYISGRGSQWAGNQQLLADEPAFAAAVAELEPFVEGQAGSLHVDVLANGE-LVIGIE 689
 QY 60 YTOPAILATSVAIYRLLEQKGYQDMVAGLSGLGYSALVAGSALDGFEDAVLVAKRGAYM 119
 Db 690 QIQLGLIGMQLTUTELWRSYGVQPDVLVGHSMGEVAAVAVAGALTPEAGLRVTATRARLM 749
 QY 120 EEARPADSGWAVLNTFVEVIEACOKASELGVVTPANTYNTPAQIVIAGEVAVDRAVE 179
 Db 750 ---APLSGGQGMALLDAPTEAL---IADFPQVTIGIYNSPQTVIAGTQID---E 800
 QY 180 LLOBAGAKRLIPLKVS---GPFHTAL--LEPASOKLAETLAQVSFSDFT--CPLVG--NT 230
 Db 801 LIARVRAQNRFPASRVNIEVAPHNPAMDALQAPMES-----ELADLTPTPTIGII 852
 QY 231 EAAVQKEDI--AQLLTQVKEPVRFYESI---GVMOEAGISNFIETIGPKVLSGFVKI 285
 Db 853 TYADLHTQIPDAEHWATNMNRPVRFQCAIASAGSGADGAYHTFIEISAPHLT---QAI 909
 QY 286 DQTAHLA 292
 Db 910 IDTLHSA 916

Search completed: June 11, 2003, 20:03:40
 Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 11, 2003, 19:58:52 ; Search time 59 Seconds
(without alignments)
1068.651 Million cell updates/sec

Title: US-09-308-397-2
Perfect score: 1518
Sequence: 1 MTKTAFIFAGGAGYLGNGR.....QTAHLAVEDQASLVALLX 306

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1518	100.0	306	Q9FBC4	Q9fbc4 streptococc
2	1048	69.0	308	16 Q9CHP8	Q9chf8 lactococcus
3	1036	68.2	312	16 Q99YD5	Q99yds streptococc
4	775.5	51.1	309	16 Q9R9V9	Q9r9v9 theamoanaer
5	713.5	47.0	308	16 Q97DA5	Q97das clostridium
6	703.5	46.3	314	16 Q9XLD2	Q9xld2 clostridium
7	657.5	43.3	312	16 Q9KQK6	Q9kqk6 vibrio chol
8	654	43.1	313	16 Q9KA02	Q9ka02 bacillus ha
9	640.5	42.2	309	16 Q9X817	Q9x817 escherichia
10	628.5	41.4	309	16 Q9ZFF6	Q9zff6 yersinia pe
11	626	41.2	312	16 Q9CJ57	Q9cjs7 pasteurella
12	612.5	40.3	309	16 Q9Z7J5	Q9z7j5 salmonella
13	611.5	40.3	307	2 Q9ZAK4	Q9zak4 vibrio mari
14	598	39.4	309	16 Q9Y0J3	Q9y0j3 ralsconia s
15	594.5	39.2	326	16 Q67041	Q67041 aquifex aeo
16	588.5	38.8	313	16 Q9Y689	Q9y689 listeria mo

17	586	38.6	313	16 Q92AK0	Q92ak0 listeria in
18	570	37.5	299	16 Q8RGX6	Q8rgx6 fusobacteri
19	570	37.5	308	16 Q9JWS8	Q9jws8 naissieriam
20	566	37.3	308	16 Q9JXR4	Q9jxr4 naissieriam
21	556.5	36.7	312	16 Q54437	Q54437 pseudomonas
22	551.5	36.3	308	16 Q99UN8	Q99un8 staphylococ
23	550.5	36.3	308	2 Q93QD4	Q93qd4 staphylococ
24	533.5	35.1	317	16 Q9PF17	Q9pf17 xyella fas
25	531	35.0	308	16 Q9JSG5	Q9jsg5 chlamydia p
26	529	34.8	305	16 Q9RT24	Q9rt24 deinococcus
27	528	34.8	308	16 Q9Z8P1	Q9z8p1 chlamydia p
28	526	34.7	293	16 Q9WZ05	Q9wzq5 thermotoga
29	503.5	33.2	400	2 Q45564	Q45564 bacillus su
30	492	32.4	331	16 Q8UGE4	Q8uge4 agrobacteri
31	485	31.9	314	16 Q9YFP2	Q9yfp2 bruceella me
32	482	31.8	314	16 Q9KGI1	Q9kgt1 rhizobium m
33	479.5	31.6	400	2 Q9J157	Q9j157 bacillus su
34	479	31.6	292	16 Q8Z061	Q8z061 anabaena sp
35	479	31.6	314	16 Q9A7P6	Q9a7p6 caulobacter
36	471.5	31.1	1110	2 Q8RL73	Q8rl73 pseudomonas
37	468	30.8	314	16 Q92GK7	Q92gk7 rickettsia
38	464.5	30.6	308	16 Q9PKF6	Q9pkf6 chlamydia m
39	461.5	30.4	306	16 Q9PJ11	Q9pj11 campylobact
40	461	30.4	315	16 Q984T5	Q984t5 rhizobium l
41	457.5	30.1	400	2 Q9R9J2	Q9r9j2 bacillus su
42	451.5	29.7	308	16 Q84241	Q84241 chlamydia t
43	445	29.3	319	16 Q9ZCJ6	Q9zcj6 rickettsia
44	440.5	29.0	349	10 Q9XGMO	Q9xgm0 brassica na
45	439.5	29.0	293	10 Q8RU07	Q8ru07 arabidopsis

ALIGNMENTS

RESULT 1
Q9FBC4 PRELIMINARY; PRT; 306 AA.
ID Q9FBC4
AC Q9FBC4;
DT 01-MAR-2001 (TREMREL. 16, Created)
DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Malonyl-CoA:ACP transacylase (Malonyl CoA-acyl carrier protein transacylase).
DE FAD OR SP0420.
GN Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R6;
RX MEDLINE=20365714; PubMed=10910344;
RA Heath R.J., Rock C.O.;
RT "A tetracycline-resistant bacterial enzyme."
RL Nature 406:145-146(2000).
RN [2]
RP SEQUENCE FROM N.A.

```

DR TIGR; SPO420; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; FabD.
DR Pfam; PF00698; Acyl_transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
KW Complete proteome.
SQ SEQUENCE 306 AA; 33091 MW; 83324189A89C2464 CRC64;

Query Match 100.0%; Score 1518; DB 16; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.9e-99;
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
DB 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
QY 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
DB 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
QY 121 EAPADSGKQWAVLNTPEVIEEACQKASEL--GVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
DB 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
QY 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
DB 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
QY 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
DB 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
QY 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 301 VALLEK 306
DB 301 VALLEK 306

RESULT 2
Q9CHF8 PRELIMINARY; PRT; 308 AA.
AC Q9CHF8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative malonyl CoA-acyl carrier protein transacylase
DE (EC 2.3.1.39).
GN FABD OR LI0773.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RA "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006311; AAK04871.1; -.
DR HSP; P25715; 1MLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; FabD.
DR Pfam; PF00698; Acyl_transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
KW Complete proteome.
SQ SEQUENCE 308 AA; 33541 MW; BAB129A7A9EE4CF CRC64;

Query Match 69.0%; Score 1048; DB 16; Length 308;
Best Local Similarity 67.5%; Pred. No. 3e-66;
Matches 308; Conservative 40; Mismatches 58; Indels 2; Gaps 1;

QY 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
DB 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
QY 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
DB 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
QY 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
DB 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
QY 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
DB 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
QY 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 301 VALLEK 306
DB 301 VALLEK 306

RESULT 3
Q99YD5 PRELIMINARY; PRT; 312 AA.
AC Q99YD5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative malonyl CoA-acyl carrier protein transacylase
DE (EC 2.3.1.39).
GN FABD OR SPV1750.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Perretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006603; AAK34494.1; -.
DR HSP; P25715; 1MLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000315; BPD_transp.
DR InterPro; IPR004410; FabD.
DR Pfam; PF00698; Acyl_transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN 1.
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 312 AA; 34214 MW; OBAC6A6D9199D5B CRC64;

Query Match 68.2%; Score 1036; DB 16; Length 312;
Best Local Similarity 68.3%; Pred. No. 2.2e-65;
Matches 209; Conservative 39; Mismatches 58; Indels 0; Gaps 0;

QY 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
DB 1 MTKTAFAGQAGQYLGMRGDFDYQYPIVKETIDRASQVLGYDLRLYLTDEEDKLNQTRY 60
QY 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
DB 61 TQPAILATSVAIYRLLOEKGYQPDWVAGSLGYSALVAGSALDFFDAVALVAKRGAYME 120
QY 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
DB 121 EAPADSGKQWAVLNTPEVIEEACQKASELGVVTPANTYNTPAQIVIAGEVAVDRAVEL 180
QY 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
DB 179 ELLOEAGAKRLIPLKVGSPFHTALLESASOKLAETLAQVSFSDFTCPVLGVNTEAAVMQKEDI 240
QY 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 239 DIAQLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
DB 241 AQLLTRQVKEPVRFYESIGVQWQAGISNFIETGPGKVLGSGFVKIKIDQTAHLAHVEDQASL 300
QY 301 VALLEK 306
DB 301 VALLEK 306

```

```
Db 121 EAPQSGKQWAVMNTDVQVIEVCQIAAKHGVPANVNTPSQIVIGQTDVAVVEL 180
Qy 181 LQAGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVSPDFTCLPVGNTAAVNMKEDI 240
Db 181 LKRGVKKRIPLNVSGPFHTALLEPASRLLAELERYNFSDPKIPLVGNTEANIMEKDI 240
Qy 241 AQLTRQVKEPVRFYFESIGVMQEGAGISNFIEIGPKVLGFGFKIDQTAHLAHVEDQASL 300
Db 241 PELLARQVNEPRFVDSVATLVESGITQFIEVGPKVLGTGFKIDKLLCTSVENWVSL 300
Qy 301 VALLEK 306
Db 301 RLFLDR 306

RESULT 4
Q8R9V9 PRELIMINARY; PRT; 309 AA.
AC Q8R9V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE (acyl-carrier-protein) S-malonyltransferase.
GN FABD OR TTE1473.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4T / JCM11007;
RX MEDLINE=2192816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AB013105; AAM24695.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 309 AA; 33891 MW; 81557C60A4AC7983 CRC64;

Query Match 51.1%; Score 775.5; DB 16; Length 309;
Best Local Similarity 50.8%; Pred. No. 5.4e-47;
Matches 154; Conservative 58; Mismatches 90; Indels 1; Gaps 1;

Qy 3 KTAFLPAGGAGYGLMGGRDFYDQYPIVKETIDRASQVLGYDLRYL-IDTEEDKLNQTRYT 61
Db 2 KTAFLPAGGAGYAGKGIYEKYEAKEIFERADEALGFNISKLCFSGPBEELMKTENT 61
Qy 62 QPAILATSVATYRLLEQKGYQPMVAGSLGYSALVAGSALDGFEDAVLVAKRGAYMEE 121
Db 62 QPAILTVSVALTRVLQKRGKPDVTAGLSLGEYSLSLVLAELDFEDAVLVKRGKYMQE 121
Qy 122 AAPADSGKQWAVLNTPEVIEEACOKASGLVVTNPANTPAQIVIAGEVAVDRAVELL 181
Db 122 VVPEGVGTMAALIGLPNEVEIEICIASVGVVEPANTNCPQLVVSQEVKAVRAVELA 181
Qy 182 QPAGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVSPDFTCLPVGNTAAVNMKEDIA 241
Db 182 KERGAKAVVLAVSAVPPHCSMLKAGELLAKELDKLVKVPVISNVTADYVQKQVK 241
Qy 242 QLLTRQVKEPVRFYFESIGVMQEGAGISNFIEIGPKVLGFGFKIDQTAHLAHVEDQASL 301
Db 242 ELLIKQVSHVLEWQSVRKWIEDGVDTFTFIEIGFKTLGSGFVKIDRSTVILNFEDESLM 301
Qy 302 ALL 304
Db 302 KAL 304

RESULT 5
Q97DAS PRELIMINARY; PRT; 308 AA.
ID Q97DAS
Q97DAS
```

```
AC Q97DAS;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase.
GN CAC3575.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Ghu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007854; AAK81498.1; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; FabD.
DR Pfam; PF00698; Acyl transf. 1.
DR TIGRFAMs; TIGR00128; fabD; 1.
KW Complete proteome.
SQ SEQUENCE 308 AA; 33811 MW; 26901C5C584AD4E2 CRC64;

Query Match 47.0%; Score 713.5; DB 16; Length 308;
Best Local Similarity 48.5%; Pred. No. 1.3e-42;
Matches 147; Conservative 55; Mismatches 97; Indels 3; Gaps 3;

Qy 1 MTKTAPLAFAGAGYGLMGGRDFYDQYPIVKETIDRASQVLGYDLRYL-IDTEEDKLNQTR 59
Db 1 MGKIAFVFGQSQYGVGMKLDYDNVQSAKETDKADEVLGFKISLCEPKGKEELNLT 60
Qy 60 YTPAILATSVATYRLLEQKGYQPMVAGSLGYSALVAGSALDGFEDAVLVAKRGAY 118
Db 61 NTQPAVLITSIAALRALEBEKGIKPDVAVAGLSLGEYSALVAGSALDGFEDAVLVKRGY 120
Qy 119 MEAAAPADSGKQWAVLNTPEVIEEACOKASELGVVTNPANTPAQIVIAGEVAVDRAV 178
Db 121 MQSAVPKIGTMAITIGLEGDVVRGICAEOSKEGIVEVANVNCPGQIVIAGEVAVESAC 180
Qy 179 ELLQEGAKRLIPLKVSQPFHTALLEPASQKLAETLAQVSPDFTCLPVGN-TEAAVMQK 237
Db 181 TKLKGARTVMSVSGPHTSMLKSAEKLKEELKNINIEDMKVPVITNVTGVDYEDK 240
Qy 238 EDIAQLLTROKQEPVRFYFESIGVMQEGAGISNFIEIGPKVLGFGFKIDQTAHLAHVEDQ 297
Db 241 DOIKGLLKKQWSSVRWEDTIRMDGDGVDTFTFIEIGFKTLSSFIKKIKRKMIFNIEKA 300
Qy 298 ASL 300
Db 301 EDL 303

RESULT 6
Q8XLH2 PRELIMINARY; PRT; 314 AA.
ID Q8XLH2
Q8XLH2
AC Q8XLH2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase.
GN FABD OR CPE1069.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
RL EMBL; AF003189; BAB0775.1; -.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; Fabd.
DR Pfam; PF00698; Acyl transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
KW Complete proteome.
SQ SEQUENCE 314 AA; 34273 MW; 801E151B80390156 CRC64;

Query Match 46.3%; Score 703.5; DB 16; Length 314;
Best Local Similarity 47.7%; Pred. No. 6.8e-42;
Matches 147; Conservative 61; Mismatches 97; Indels 3; Gaps 3;

QY 1 MTKTAFLPAGOGAOLVGMGRDFDYQYPIVKTIDRASOVLYGLYLDTE-EDKLNQTR 59
DB 1 MAKLGFLPAGOGAOLVGMGRDFDYQYPIVKTIDRASOVLYGLYLDTE-EDKLNQTR 60

QY 60 YTPAILATSVAIYRLLOEKYQPDVMVAGLSGLSEYLSALVAGSGALDFEDAVLVAKRGAYM 119
DB 61 FTQPAIITTNMAILLTALDKLVKSHISGSLSGEYLSALIHSGAINPFDGVLVKKRGKFM 120

QY 120 BEAPAPSGKMWAVLNTPEVIEEACOKASELGWVTPANTYNTPAQIVIAGEVAVDRAVE 179
DB 121 QEVAEGTGGWAVLNTPEVIEEACOKASELGWVTPANTYNTPAQIVIAGEVAVDRAVE 180

QY 180 LLOBAGAKRLIPLKVSPPFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEA-AVMQKE 238
DB 191 FIKEVGG-RAIKLPVSAPFFHSCMLQPAEKLDELINISINKLNGIWMNVKGEVLEDD 239

QY 239 DIAQLLTQVKEPVRFYESIGVMOEAGISNFIETGPGKVLGSGFVKKIDQTAHLAHVEDQA 298
DB 240 NIELLTQVKKPVLFIENDIEKMTESGVDTEIETGPGKALSGFVKKINKVTVLNVEDLK 299

QY 299 SLVALLEK 306
DB 300 SLEKTLK 307

RESULT 7
Q9KQH6 PRELIMINARY; PRT; 312 AA.
AC Q9KQH6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase.
GN VC2022.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
RL EMBL; AF004276; AAF95170.1; -.
DR HSSP; P25715; 1MLA.
DR TIGR; VC2022; -.

```

```

DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; Fabd.
DR Pfam; PF00698; Acyl transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
KW Complete proteome.
SQ SEQUENCE 312 AA; 33217 MW; 458A38C7E1D2356F CRC64;

Query Match 43.3%; Score 657.5; DB 16; Length 312;
Best Local Similarity 47.4%; Pred. No. 1.2e-38;
Matches 146; Conservative 58; Mismatches 97; Indels 7; Gaps 4;

QY 1 MTKTAFLPAGOGAOLVGMGRDFDYQYPIVKTIDRASOVLYGLYLDTE-EDKLNQTR 59
DB 6 MSKFAIVEFGGSGAVGMLDLAEQYAVVKTFAEASEVLGYDLWALVQDGPVEDLNQTF 65

QY 60 YTPAILATSVAIYRLLOEKYQPDVMVAGLSGLSEYLSALVAGSGALDFEDAVLVAKRGAY 118
DB 66 RTQPALLAASVAIWRVMOQLGLEQPAVLGHSGLSEYLSALVAGVIDFKQAILVELRQOL 125

QY 119 MEERAPADSGKMWAVLNTPEVIEEACOKASELGWVTPANTYNTPAQIVIAGEVAVDRAV 178
DB 126 WOQAVPAGTGMAYALIGLEDAIAKACADAAGQGVSPVNFNSGQVVIAGKDAVERAG 185

QY 179 ELLQAGAKRLIPLKVSPPFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEA-AVMQKE 238
DB 186 VLCKEAGAKRALPLPSVSPSHCALMKPAADBLAKTLAELEFNAPQIPVNNVD--VVAET 243

QY 239 D---IAQLLTQVKEPVRFYESIGVMOEAGISNFIETGPGKVLGSGFVKKIDQTAHLAHE 295
DB 244 DVKTKDALIKQLVSPVRWTECVQMSAQGVKEKLIEMGFKVLTLTKRIVTKTLEGVAVN 303

QY 296 DOASLVAL 303
DB 304 DVASLDAV 311

RESULT 8
Q9KA02 PRELIMINARY; PRT; 313 AA.
AC Q9KA02;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39).
GN FABD OR BH2492.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001515; BAC06211.1; -.
DR HSSP; P25715; 1MLA.
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000429; DEAD box.
DR InterPro; IPR004410; Fabd_box.
DR Pfam; PF00698; Acyl transf; 1.
DR TIGRFAMs; TIGR00128; fabd; 1.
DR PROSITE; PS00039; DEAD ATP HELICASE; UNKNOWN_1.
KW Transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 313 AA; 33585 MW; B009ACA5AD8C3968 CRC64;

Query Match 43.1%; Score 654; DB 16; Length 313;
Best Local Similarity 45.6%; Pred. No. 2.1e-38;
Matches 141; Conservative 58; Mismatches 106; Indels 4; Gaps 4;

```


Qy	1	MTKTAFLPAGQAQVILGMRDFFYDOYPIVKSTIDRASQVILGYDL-RVLIIDTEEDKLNQTR	59
Db	1	MAKVAFLPFGQSQSGVNGSELLSBEK-AKEIFDEADERLGYSLSSIMFGEPEKURRTE	59
Qy	60	YTQPAILATSVAIYRILLOEKGVQPDMAVAGLSGEYSALVAGGALDPEDDAVALVAKRGAYM	119
Db	60	NTQPALLTMTSTAVLSLVREYGIKPDYTAGHLSGEYSALVAGSULTFADAVAYVHHRGLFM	119
Qy	120	EEAAPADSGKMVAVLNTPVEIEEACQKASELG-VVTPNANTPAQIVTIAGEVAVVDRAV	178
Db	120	EEAVPFGEGAMAAIILGMRDELEQVTKRVTEAGAVWELANLPCPGQIVISGSAGEVEQAS	179
Qy	179	ELLQEAGAKGLIPLKVSQGFHTALLEPASOKLATIAQVSEFDTCPVLGNTESAAMQK-	237
Db	180	EEAKEAGAKRVIPLOVSGFPFHSSLMKPAEKLAVLADLAIAOAAPEVIANVTADIVQKA	239
Qy	238	EDIAQLLTQVKEPVRVFEIGWQEAGISNFIEIGPGKVLSGFVKKIDOTAHLAHVQD	297
Db	240	ADIRASLIEQVYSPVRWEDTVRMLELGVDTFVEIGSVLGLVRKVQRVNVVFSVSDR	299
Qy	298	ASLVALLEK 306	
Db	300	ASIEAMVKK 308	

```

RESULT 9
Q8X817
ID QX817 PRELIMINARY; PRT; 309 AA.
AC QX817;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Malonyl-CoA [acyl-carrier-protein] transacylase.
GN FABD OR Z1731 OR ECS1470.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
CX NCBI_TaxID=83334;
CN [1]_TaxID=83334;
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J.J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimailanta E.T., Potancusis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533 (2001).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohsesu E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhrata S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res 8:11-22 (2003).
DR ENBL; A5C05319; BAG5868.1; -
DR ENBL; AF002555; BAB34893.1; -
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR004410; FabD.
DR Pfam; PF00698; Acyl transf; 1.
DR TIGRFAMs; TIGR00128; fabD; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 32412 MW; 3DCB41120D526F71 CRC64;

```

Matches	141;	Conservative	59;	Mismatches	103;	Indels	7;	Gaps	4;
Qy	1	MTKTAFLPAGGAQYLGMRGDFVDQYPIVKEITIDRASQVLYDRLYLIDTEE---	DKLNQ	57					
Db	1	MTQFAFVFGQGSQTVMGLADMAASYPVIEETFAEASAAALGYDLWAL--	TQOGPAEELNK	58					
Qy	58	TRYTQFAILATSVAIYRLLEQKGY-QPDMVAGLSLGEYSALVASGALDPEDAVALKVRG	116						
Db	59	TWOTQPALITASVALYRVWQHGKAPAMVAGHSLGEYSALVCAGVIDFADAVRLVEMRG	118						
Qy	117	AYNEEAAPADSGKMWAVLNTPEVIEBEACQKASELGVVTPANVNTPAQIVIAGEVAVDR	176						
Db	119	KFMQEAAPVPGTGMATAIIGLDDASIGKACEEAAGQVSPVNFNSPGQVVIAGHKEAVER	178						
Qy	177	AVELLQEAAGAKRLIPKLVSGPPHTALLESAPASKLAETLQAQVSFSDFTCPVLVGNTEAAVMQ	236						
Db	179	AGAACAKAAGAKRALPLFVSPVSHCALMKPAADKLAVELAKIFENATFVPVNVNDVKCET	238						
Qy	237	KED-IAQLITRQVKEPRVYESIGNVQEAAGISNFIIGPGKVLVSFGVKKIDQTAHLAHVE	295						
Db	239	NGDAIRDALVRQLYNPVQWTKSVYVMAAOGVEHLXEVGPGKVLGTGKTRIVDTLTASALN	298						
Qy	296	DQASLVALLE 305							
Db	299	EPSMAAALE 308							
RESULT 10									
Q8ZFT6	ID	Q8ZFT6	PRELIMINARY;	PRT;	309	AA.			
AC	Q8ZFT6;								
DT	01-MAR-2002	(TEMBLrel. 20, Created)							
DT	01-MAR-2002	(TEMBLrel. 20, Last sequence update)							
DT	01-JUN-2002	(TEMBLrel. 21, Last annotation update)							
DE	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39).								
GN	FABD OR YPO1598.								
OS	Yersinia pestis.								
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;								
CC	Yersinia.								
OX	NCBI_TaxID=632;								
RN	[1]								
RN	SEQUENCE FROM N.A.								
RP	STRAIN=CO-92 / BIOVAR ORIENTALIS;								
RC	MEDLINE=21470413; PubMed=11586360;								
RX	Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,								
RA	Prentice M.B., Sebainia D., James K.D., Churcher C., Mungall K.L.,								
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,								
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,								
RA	Feltwell T., Hamlin N., Holtroyd S., Jagels K., Kariyshev A.V.,								
RA	Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,								
RA	Simmonds M., Skellon J., Stevens K., Whitehead S., Bartell B.G.,								
RT	"genome sequence of Yersinia pestis, the causative agent of plague.";								
RL	Nature 413:523-527(2001).								
RL	EMBL; AJ414149; CAC90420.1; -								
DR	InterPro; IPR001227; Ac transferase.								
DR	InterPro; IPR004410; FabD.								
DR	Pfam; PF00698; Acyl transf. 1.								
DR	TIGRfams; TIGR00128; fabd; 1.								
KV	Transferase; Acyltransferase; Complete proteome.								
SQ	SEQUENCE 309 AA; 32651 MW; 5F090049FF9A848B CRC64;								
Query Match 41.4%; Score 628.5; DB 16; Length 309;									
Best Local Similarity 45.5%; Pred. No. 1.3e-36;									
Matches	140;	Conservative	56;	Mismatches	109;	Indels	3;	Gaps	3;
Qy	1	MTKTAFLPAGGAQYLGMRGDFVDQYPIVKEITIDRASQVLYDRLYLIDT-EEDKLNQTR	59						
Db	1	MSKFAVFPQGSQSLMGLADLAQAQPIVETATSEASSVLGYDLWGLVQOGPAEELNKTW	60						
Qy	60	YTQFAILATSVAIYRLLEQKGY-QPDMVAGLSLGEYSALVASGALDPEDAVALKVRGAY	118						
Db	61	OTQPALITASVAIYRLLEQKGY-QPDMVAGLSLGEYSALVASGALDPEDAVALKVRGAY	118						
Db	61	OTQPALITASVAIYRLLEQKGY-QPDMVAGLSLGEYSALVASGALDPEDAVALKVRGAY	118						

```

QY 119 MEERAPADSGKQVAVLNTPTVEIEBACQKASSELGVVTPANNYTPAQIVIAGEVAVDRAV 178
DB 121 MQEAVPEGTGAMYAIIGLDNESIAKACESAQGVVSPVNFSPGVVIAGNKDAVERAG 180
QY 179 ELLOQAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEA-AVMQK 237
DB 181 AACKAAGAKRALPLVPSVPSHCALMKPADKLAVALKIDFQEPFLFPVNNVDVKAESP 240
QY 238 EDIAQLLTQVKEPVRFYESIGVMOEAGISNFIIEIGPKVLSGFVKKIDQTAHLAHVEDQ 297
DB 241 VAIRGALVRLQYNPVRWTESVEFFIAAGVELLLEIGPKVLTGLTKRIVDSLAATAVNDV 300
QY 298 ASLVALLE 305
DB 301 VTLNSALE 308

RESULT 11
Q9CJS7 PRELIMINARY; PRT; 312 AA.
AC Q9CJS7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE FabD.
GN FADB OR PM1915.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Fm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AB006228; AAK03999.1; -.
DR HSSP; P25715; IMLA.
DR InterPro; IPR001227; AC transferase.
DR InterPro; IPR004410; FADB.
DR Pfam; PF00698; Acyl transferf. 1.
DR TIGSFams; TIGR00128; fadb; 1.
KW Complete proteome.
SQ SEQUENCE 312 AA; 33198 MW; 36E1A83C0A26CD89 CRC64;

Query Match 41.2%; Score 626; DB 16; Length 312;
Best Local Similarity 44.7%; Pred. No. 2e-36;
Matches 138; Conservative 57; Mismatches 110; Indels 4; Gaps 3;

QY 1 MTKTAPLFAQGAQVLMGRDFYDQYPIVKETIDRASQVLYGVDLRYLIDT-EEDKLNQTR 59
DB 1 MKKFAMVFPQGSQAVGMLAELATSPYVVEETFKQASDVLGYDLWQLVQOGPFAEELNKTW 60
QY 60 YTPQAILATSVAIYRLLOEK--GYQDMVAGLSGEYSALVASGALDPEDAVALVAKRGA 117
DB 61 QTQPALLAASVAIYRWQEKYHLPKPDVWAGHSLGEYSALVCAGALDFQDAVLVELRGK 120
QY 118 YNEEAPADSGKQVAVLNTPTVEIEBACQKASSELGVVTPANNYTPAQIVIAGEVAVDRA 177
DB 121 LMQAVPEGTGAMYAIIGLDNEAIIISACADAQGEVWSAVNFSPGVVIAAGAAVERA 180
QY 178 VELLOQAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGNTEAAV-MQ 236
DB 181 AAACKDAGAKRALPLAVSPVSHCALMKPADQCLAVSLNDIAIRPTTAVINNVDVACETE 240
QY 237 KEDIAQLLTQVKEPVRFYESIGVMOEAGISNFIIEIGPKVLSGFVKKIDQTAHLAHVED 296
DB 241 NSEIRHALVRLQYSPVRWTETVERMAKQGVQLVEVGNKVLTLGLTKRIVADLQATAVND 300
QY 297 QASLVALLE 305
DB 301 LTLNSAVDE 309

```

RESULT 12

```

Q827J5 PRELIMINARY; PRT; 309 AA.
ID Q827J5;
AC Q827J5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Malonyl CoA-acyl carrier protein transacylase.
GN STY1233.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL; AL627269; CAD08318.1; -.
DR InterPro; IPR001227; AC transferase.
DR InterPro; IPR004410; FADB.
DR Pfam; PF00698; Acyl transferf. 1.
DR TIGSFams; TIGR00128; fadb; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 32435 MW; 83B4B7FE13C8CB3 CRC64;

Query Match 40.3%; Score 612.5; DB 16; Length 309;
Best Local Similarity 43.9%; Pred. No. 1.8e-35;
Matches 138; Conservative 62; Mismatches 97; Indels 17; Gaps 5;

QY 1 MTKTAPLFAQGAQVLMGRDFYDQYPIVKETIDRASQVLYGVDLRYLIDT-EEDKLNQ 57
DB 1 MTQFAPVFPQGSQSVGMLAEMANYPVVEETFAEASALGYDLNAL--TQOGPAEELNK 58
QY 58 TRYTPAILATSVAIYRLLOEK-GYQDMVAGLSGEYSALVASGALDPEDAVALVAKRG 116
DB 59 TWQTQALLTASVALVWRVWQOGKMPALMAGHSLGEYSALVCAGVINEADAVRLVEMRG 118
QY 117 AYNEEAPADSGKQVAVLNTPTVEIEBACQKASSELGVVTPANNYTPAQIVIAGEVAVDRA 176
DB 119 KFMQEAIVPEGTGMSAIIIGLDASIAKACEEAGQVSPVNFSPGVVIAAGHKEAVER 178
QY 177 AVELLOQAGAKRLIPLKVSQGFHTALLEPASQKLAETLAQVSPSDFTCPLVGN-----T 230
DB 179 AGACACAAKAKRALPLVPSVPSHCALMKPADKLAVALAKITFSAPTVPVNNVDVKCET 238
QY 231 EAAVMKEDIQALLTQVKEPVRFYESIGVMOEAGISNFIIEIGPKVLSGFVKKIDQTAH 290
DB 239 DAAA-----IRDALVRLQYNPQVTKSVSEFFIAAQGVHELYEYVGPGLTKRIVDTLT 293
QY 291 LAHVEDQASLVALL 304
DB 294 ASALNEPAALSAL 307

RESULT 13
Q9RA34 PRELIMINARY; PRT; 307 AA.
ID Q9RA34;
AC Q9RA34;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

```

